

STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

ATTORNEYS AT LAW

1100 NEW YORK AVENUE, N.W., SUITE 600

WASHINGTON, D.C. 20005-3934

www.skgf.com

PHONE: (202) 371-2600 FACSIMILE: (202) 371-2540

ROBERT GREENE STERNE  
EDWARD J. KESSLER  
JORGE A. GOLDSTEIN  
SAMUEL L. FOX  
DAVID K.S. CORNWELL  
ROBERT W. ESMOND  
TRACY-GENE G. DURKIN  
MICHELE A. CIMBALA  
MICHAEL B. RAY  
ROBERT E. SOKOHL  
ERIC K. STEFFE  
MICHAEL Q. LEE

STEVEN R. LUDWIG  
JOHN M. COVERT\*  
LINDA E. ALCORN  
RAZ E. FLESHNER  
ROBERT C. MILLONIG  
MICHAEL V. MESSINGER  
JUDITH U. KIM  
TIMOTHY J. SHEA, JR.  
DONALD R. MCPHAIL  
PATRICK E. GARRETT  
STEPHEN G. WHITESIDE  
JEFFREY T. HELVEY\*

HEIDI L. KRAUS  
JEFFREY R. KURIN  
RAYMOND MILLIEN  
PATRICK D. O'BRIEN  
LAWRENCE B. BUGAISKY  
CRYSTAL D. SAYLES\*  
EDWARD W. YEE  
ALBERT L. FERRO\*  
DONALD R. BANOWIT  
PETER A. JACKMAN  
MOLLY A. MCCALL  
TERESA U. MEDLER

JEFFREY S. WEAVER  
KRISTIN K. VIDOVICH  
KENDRICK P. PATTERSON  
DONALD J. FEATHERSTONE  
GRANT E. REED  
VINCENT L. CAPUANO  
JOHN A. HARROUN\*  
MATTHEW M. CATLETT\*  
NATHAN K. KELLEY\*  
ALBERT J. FASULO II\*  
W. BRIAN EDGE\*

KAREN R. MARKOWICZ  
SUZANNE E. ZISKA\*\*  
BRIAN J. DEL BUONO\*\*  
ANDREA J. KAMAGE\*\*  
NANCY J. LEITH\*\*  
TARJA H. NAUKKARINEN\*\*

\*BAR OTHER THAN D.C.

\*\*REGISTERED PATENT AGENTS

April 26, 2000

WRITER'S DIRECT NUMBER:

(202) 789-5519

INTERNET ADDRESS:

KVIDOVIC@SKGF.COM

Commissioner for Patents  
Washington, D.C. 20231

*Box Patent Application*

Re: U.S. Non-Provisional Continuation Patent Application under 37 C.F.R. § 1.53(b)  
Appl. No. To be assigned; Filed: Herewith  
For: **Mutant DNA Polymerases and Uses Thereof**  
Inventor: Chatterjee, D. K.  
Our Ref: 0942.3600003/RWE/RCM

Sir:

The following documents are forwarded herewith for appropriate action by the U.S.  
Patent and Trademark Office:

1. USPTO Utility Patent Application Transmittal Form PTO/SB/05;
2. U.S. Utility Patent Application entitled:

Mutant DNA Polymerases and Uses Thereof and naming as inventor(s):

Deb K. Chatterjee

the application consisting of:

- a. A specification containing:

Commissioner for Patents

April 26, 2000

Page 2

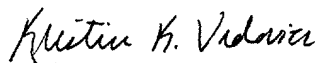
- i. 30 pages of description prior to the claims;
- ii. 5 pages of claims (28 claims);
- iii. a one (1) page abstract;
- iv. 19 pages of a paper copy of the sequence listing;
- b. 9 sheets of drawings: (Figures 1-4, 5A, 5B, and 6-7);
3. Copy of original Declaration executed by the inventor;
4. Fee Transmittal Form (*in duplicate*);
5. Our check no. 27359 in the amount of \$690.00 to cover the basic filing fee;
6. Authorization to Treat a Reply As Incorporating An Extension of Time Under 37 C.F.R. § 1.136(a)(3) (*in duplicate*); and
7. Two (2) return postcards.

It is respectfully requested that, of the two attached postcards, one be stamped with the filing date of these documents and returned to our courier, and the other, prepaid postcard, be stamped with the filing date and unofficial application number and returned as soon as possible.

This patent application is being submitted under 37 C.F.R. § 1.53(b).

Respectfully submitted,

STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.



Kristin K. Vidovich  
Attorney for Applicant  
Registration No. 41,448

RWE/KKV:amm

Enclosures

P:\USERS\kvidovich\0942\3600003\pto trl con  
SKGF Rev 4/5/00 mac

009240 76426950

Please type a sign(+) inside this box



PTO/SB/05 (2/98)

Approved for use through 09/30/2000 OMB 0651-0032  
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

<b>UTILITY PATENT APPLICATION TRANSMITTAL</b> <i>(Only for new nonprovisional applications under 37 CFR § 1.53(b))</i>	Attorney Docket No.	0942.3600003/RWE/RCM
	First Inventor or Application Identifier	Deb K. CHATERJEE
	Title	Mutant DNA Polymerases and Uses Thereof
	Express Mail Label No.	

<b>APPLICATION ELEMENTS</b> <i>See MPEP chapter 600 concerning utility patent application contents</i>	ADDRESS TO: Assistant Commissioner for Patents Box Patent Application Washington, DC 20231
---	--

1. ☒ \* Fee Transmittal Form (e.g., PTO/SB/17)  
*(Submit an original, and a duplicate for fee processing)*

2. ☒ Specification [Total Pages 55 ]  
*(preferred arrangement set forth below)*  
- Descriptive title of the Invention  
- Cross References to Related Applications  
- Statement Regarding Fed sponsored R & D  
- Reference to Microfiche Appendix  
- Background of the Invention  
- Brief Summary of the Invention  
- Brief Description of the Drawings (if filed)  
- Detailed Description  
- Claim(s)  
- Abstract of the Disclosure

3. ☒ Drawing(s) (35 U.S.C. 113) [Total Sheets 9 ]

4. ☒ Oath or Declaration [Total Pages 2 ]

a. ☐ Newly executed (original or copy)

b. ☒ Copy from a prior application (37 CFR 1.63(d)) *(for continuation/divisional with Box 17 completed)*  
[Note Box 5 below]

i. ☐ DELETION OF INVENTOR(S)  
Signed statement attached deleting inventor(s) named in the prior application, see 37 CFR §§ 1.63(d)(2) and 1.33(b).

5. ☒ Incorporation By Reference *(useable if Box 4b is checked)*  
The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 4b, is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.

6. ☐ Microfiche Computer Program *(Appendix)*

7. Nucleotide and/or Amino Acid Sequence Submission *(if applicable, all necessary)*

a. ☐ Computer Readable Copy

b. ☐ Paper Copy (identical to computer copy)

c. ☐ Statement verifying identity of above copies

#### ACCOMPANYING APPLICATION PARTS

8. ☐ Assignment Papers (cover sheet & document(s))

9. ☐ 37 CFR 3.73(b) Statement ☐ Power of Attorney  
*(when there is an assignee)*

10. ☐ English Translation Document *(if applicable)*

11. ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS Citations

12. ☐ Preliminary Amendment

13. ☒ Return Receipt Postcard (MPEP 503) (2)  
*(Should be specifically itemized)*

14. ☐ \*Small Entity Statement(s) ☐ Statement filed in prior application, Status still proper and desired  
*(PTO/SB/09-12)*

15. ☐ Certified Copy of Priority Document(s)  
*(if foreign priority is claimed)*

16. ☒ Other: 37 C.F.R. § 1.136(a)(3) Authorization

☐ Other:

\*NOTE FOR ITEMS 1 & 14: IN ORDER TO BE ENTITLED TO PAY SMALL ENTITY FEES, A SMALL ENTITY STATEMENT IS REQUIRED (37 C.F.R. § 1.27), EXCEPT IF ONE FILED IN A PRIOR APPLICATION IS RELIED UPON (37 C.F.R. § 1.28)

17. If a CONTINUING APPLICATION, check appropriate box, and supply the requisite information below and in a preliminary amendment:

☒ Continuation

☐ Divisional

☐ Continuation-in-Part (CIP) of prior application No: 08/576,759

Prior application information: Examiner Wax, R.

Group/Art Unit: 1652

#### 18. CORRESPONDENCE ADDRESS

☐ Customer Number  
or Bar Code Label

*(Insert Customer No. or Attach bar code label here)*

or ☒ Correspondence  
address below

NAME	STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.				
	Attorneys at Law				
ADDRESS	Suite 600, 1100 New York Avenue, N.W.				
CITY	Washington	STATE	DC	ZIP CODE	20005-3934
COUNTRY	USA	TELEPHONE	(202) 371-2600	FAX	(202) 371-2540

NAME (Print/Type)	Kristin K. Vidovich	Registration No. (Attorney/Agent)	41,448
SIGNATURE	<u>Kristin K. Vidovich</u>	Date	<u>Apr 12, 2000</u>

Burden Hour Statement: this form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

# Mutant DNA Polymerases and Uses Thereof

## *Cross Reference to Related Applications*

5 This application is a continuation of Serial No. 08/537,397, filed October 2, 1995, entitled Mutant DNA Polymerases and Uses Thereof, which is a continuation-in-part of Serial No. 08/525,057 of Deb K. Chatterjee, filed September 8, 1995, also entitled Mutant DNA Polymerases and the Use Thereof. The content of both of these applications is specifically incorporated herein by reference.

## *Field of the Invention*

10 This invention relates to molecular cloning and expression of mutant DNA polymerases that are particularly useful in DNA sequencing reactions.

## *Background of the Invention*

15 DNA polymerases synthesize the formation of DNA molecules from deoxynucleotide triphosphates using a complementary template DNA strand and a primer. DNA polymerases synthesize DNA in the 5'-to-3' direction by successively adding nucleotides to the free 3'-hydroxyl group of the growing strand. The template strand determines the order of addition of nucleotides *via* Watson-Crick base pairing. In cells, DNA polymerases are involved in repair synthesis and DNA replication.

20 Bacteriophage T5 induces the synthesis of its own DNA polymerase upon infection of its host, *Escherichia coli*. The T5 DNA polymerase (T5-DNAP) was purified to homogeneity by Fujimura RK & Roop BC, *J. Biol. Chem.* 25:2168-2175 (1976). T5-DNAP is a single polypeptide with a molecular weight of about 96 kilodaltons. This polymerase is highly processive and, unlike T7 DNA  
25 polymerase, does not require thioredoxin for its processivity (Das SK & Fujimura

RK, *J. Biol. Chem.* 252:8700-8707 (1977); Das SK & Fujimura RK, *J. Biol. Chem.* 254:1227-1237 (1979)).

Fujimura RK *et al.*, *J. Virol.* 53:495-500 (1985) disclosed the approximate location of the T5-DNAP gene on the physical restriction enzyme map generated by Rhoades, *J. Virol.* 43:566-573 (1982). DNA sequencing of the fragments of this corresponding region was disclosed by Leavitt & Ito, *Proc. Natl. Acad. Sci. USA* 86:4465-4469 (1989). However, the authors did not reassemble the sequenced fragments to obtain expression of the polymerase.

Copending application Serial No. 08/370,190, filed January 9, 1995, discloses a DNA polymerase from an eubacterium, *Thermotoga neapolitana* (Tne). A partial restriction map and a partial DNA sequence of this DNA polymerase gene have been established.

An oligonucleotide-directed, site-specific mutation of a T7 DNA polymerase gene was disclosed by Tabor S & Richardson CC, *J. Biol. Chem.* 264:6447-6458 (1989).

The existence of a conserved 3'-to-5' exonuclease active site present in a number of DNA polymerases is discussed in Bernard *et al.*, *Cell* 59:219-228 (1989). T5 DNA polymerase which lacks 3'-to-5' exonuclease activity is disclosed in U.S. Patent No. 5,270,179.

In molecular biology, DNA polymerases have several uses. In cloning and gene expression experiments, DNA polymerases are used to synthesize the second strand of a single-stranded circular DNA annealed to an oligonucleotide primer containing a mutated nucleotide sequence. DNA polymerases have also been used for DNA sequencing by the Sanger Dideoxy method. For example, the Klenow fragment, Taq DNA polymerase and T7 DNA polymerase lacking substantial exonuclease activity, are useful for DNA sequencing. Such DNA sequencing procedures are carried out by annealing a primer to a DNA molecule to be sequenced, incubating the annealed mixture with a DNA polymerase, and four deoxynucleotide triphosphates in four vessels each of which contains a different DNA synthesis terminating agent (e.g. a dideoxynucleoside triphosphate). The

agent terminates at a different specific nucleotide base in each of the four vessels. The DNA products of the incubating reaction are separated according to their size so that at least part of the nucleotide base sequence of the DNA molecule can be determined.

5 Residues in DNA polymerases important for binding of nucleotides have been investigated by Polesky, A.H. *et al.*, *J. Biol. Chem.* 265:14579-14591 (1990) and Astatke M *et al.*, *J. Biol. Chem.* 270:1945-1954 (1995).

While several DNA polymerases are known, there exists a need in the art for additional DNA polymerases having properties suitable for DNA synthesis, DNA sequencing, and DNA amplification.

### *Summary of the Invention*

10 The present invention helps satisfy these needs in the art of providing additional DNA polymerases and uses therefor. This invention is related to the discovery that it is possible to prepare mutant DNA polymerases that incorporate dideoxynucleotides into a synthesized DNA molecule with about the same efficiency that deoxynucleotides are incorporated. Such mutant DNA polymerases may be used to prepare sequencing ladders having bands of approximately equal intensity.

15 Thus, the present invention is related to a mutant DNA polymerase that incorporates dideoxynucleotides with about the same efficiency as deoxynucleotides, wherein the native DNA polymerase favors the incorporation of deoxynucleotides over dideoxynucleotides. Examples of the mutant DNA polymerase include a mutant Klenow fragment of DNA polymerase, *e.g.* of *E.*

20

*coli*, a mutant T5 DNA polymerase, a mutant Taq polymerase, a mutant *Thermatoga maritima* (Tma) DNA polymerase (U.S. Patent 5,374,553), and a mutant of Tne polymerase.

The invention also relates to a DNA molecule which codes for the mutant DNA polymerase of the present invention as well as host cells comprising the DNA molecule.

The invention also relates to a method for producing a protein, wherein said protein has a mutant DNA polymerase activity and incorporates dideoxynucleotides with about the same efficiency as deoxynucleotides, said method comprising the steps of:

- (i) culturing a host cell containing the DNA molecule of the invention, and
- (ii) isolating said protein from said host cell.

Examples of such mutant DNA polymerase proteins include mutant T5 DNA polymerase, wherein Tyr<sup>570</sup> is substituted for Phe<sup>570</sup> of native T5 DNA polymerase; mutant Taq DNA polymerase, wherein Tyr<sup>667</sup> is substituted for Phe<sup>667</sup> of native Taq DNA polymerase; mutant Klenow fragment DNA polymerase, wherein Tyr<sup>762</sup> is substituted for Phe<sup>762</sup> of Klenow DNA polymerase; mutant Tne DNA polymerase, wherein Tyr<sup>67</sup> is substituted for Phe<sup>67</sup> of Tne DNA polymerase, as numbered in Figure 4; and a mutant Tma DNA polymerase, wherein Tyr<sup>730</sup> is substituted for Phe<sup>730</sup>.

In addition, this invention also relates to mutant DNA polymerases, that, in addition to incorporating dideoxynucleotides into a DNA molecule about as efficiently as deoxynucleotides, has substantially reduced 5'-to-3' exonuclease activity, substantially reduced 3'-to-5' exonuclease activity, or both substantially reduced 5'-to-3'-exonuclease activity and substantially reduced 3'-to-5' exonuclease activity. By way of example, such a mutant DNA polymerase can be a T5 DNA polymerase, a Tne DNA polymerase, a Klenow fragment DNA polymerase, a Taq DNA polymerase or a Tma DNA polymerase. This invention also relates to DNA molecules coding for mutant DNA polymerases with

substantially reduced exonuclease activity, host cells comprising the DNA molecule, and methods of producing these mutant DNA polymerases.

### ***Brief Description of the Drawings***

Figure 1 is a map of the T5 DNA polymerase expression vector pSportT5#3.

Figure 2 is a map of the Taq DNA polymerase expression vector pTTQ-Taq.

Figure 3 is a restriction map of plasmids pSport-Tne and pUC-Tne. The locations of the Tne DNA polymerase, as well as the region containing the O-helix homologous sequence, are indicated.

Figure 4 depicts the nucleotide and deduced amino acid sequences, in all 3 reading frames, of the C-terminal portion, including the O helix region, of the Tne DNA polymerase gene.

Figure 5A schematically depicts the construction of plasmids pUC-Tne (3'-5') and pUC-TneFY from pUC-Tne.

Figure 5B schematically depicts the construction of plasmids pTrcTne35 and pTrcTneFY from pUCTne(3'-5') and pUC-TneFY, respectively.

Figure 6 schematically depicts the construction of pTrcTne35FY from pUC-Tne (3'-5') and pUC-TneFY.



Figure 7 schematically depicts the construction of plasmids pTTQTne535FY and pTTQTne5FY.

### *Detailed Description of the Invention*

One of the applications of DNA polymerases, particularly the *E. coli* DNA polymerase I family, is in DNA sequencing. Of the known polymerases, the large fragment (Klenow fragment) of *E. coli* DNA polymerase I, T7 DNA polymerase, and Taq DNA polymerase are used more frequently than other DNA polymerases.

The DNA polymerase of *E. coli* bacteriophage T5 has recently been cloned and expressed. See U.S. Patent Nos. 5,270,179 and 5,047,342. The T5 DNA polymerase is a highly processive polymerase and does not require any accessory protein, such as thioredoxin, to be processive. Although T5 DNA polymerase is capable sequencing DNA in the presence of dideoxynucleoside triphosphates, it requires 20-30 fold more concentrated solutions compared to the concentration for the deoxynucleotide triphosphates to generate sequencing ladders. DNA sequencing with other polymerases such as Klenow fragment and Taq DNA polymerase also requires more dideoxynucleotides, similar to T5 DNA polymerase, to generate sequencing ladders.

T7 DNA polymerase, on the other hand, requires thioredoxin for processivity and almost eqimolar or less concentrations of dideoxynucleotides to deoxynucleotides to generate suitable sequencing ladders. The most important difference in the sequencing ladder produced by T7 DNA polymerase compared to others is that it produces bands with equal intensity throughout the sequence, while Klenow fragment, T5 DNA polymerase, Tne DNA polymerase and Taq DNA polymerases produced sequence dependent uneven band intensity. Thus, T7 DNA polymerase is more non-discriminating and more efficiently incorporates dideoxynucleotides into DNA; while T5, Taq, Tne, and Tma DNA polymerase,

and Klenow fragment are more discriminating and incorporate dideoxynucleotides inefficiently.

The Tne DNA polymerase has a molecular weight of about 100 kDa. This polymerase is extremely thermostable, showing more than 50 percent activity after being heated for 60 minutes at 90°C with or without detergent. Thus, the Tne DNA polymerase is more thermostable than Taq polymerase.

The Tne DNA polymerase of the invention can be isolated from any strain of *Thermatoga neapolitana*, which produces a DNA polymerase having a molecular weight of about 100 kDa. The most preferred *Thermatoga* strain for isolating the DNA polymerase of the invention was isolated from an African continental solfataric spring (Winberger *et al.*, *Arch. Microbiol.* 151:506-512 (1989)) and may be obtained from the Deutsche Sammlung von Microorganismen und Zellkulturen GmbH, Braunschweig, Fed. Rep. Germany, as Deposit No. 5068.

The recombinant clone containing the gene encoding DNA polymerase (DH10B/pUC-Tne) was deposited on September 30, 1994, with the Patent Culture Collection, Northern Regional Research Center, USDA, 1815 N. University Street, Peoria, IL 61604, USA, as Deposit No. NRRL B-21338.

The amino acid sequence comparison of all of these DNA polymerases suggests that all contain the conserved dNTP binding amino acids. Crystal structure as well as biochemical studies suggest that several amino acids, such as Lys and Tyr, present in the O-helix are important in dNTP binding. Both of these amino acids and several other amino acids are conserved in Klenow fragment, T5, Taq, Tne and T7 DNA polymerases (Poleskey, A. H. *et. al.*, *J. Biol. Chem.* 265:14579-14591 (1990)). Thus, amino acid(s) directly or indirectly involved in dNTP binding may be responsible for discrimination of dideoxynucleotides. By incorporating active regions of T7 DNA polymerase (which do not discriminate) into other polymerases, mutant DNA polymerases were constructed, which do not discriminate against dideoxynucleotides. The invention relates to this discovery.

Amino acid residues of T5 DNA polymerase are numbered herein as numbered in U.S. patent No. 5,270,179 and Leavitt and Ito, *Proc. Natl. Acad. Sci USA* 86:4465-4469 (1989).

Amino acid residues of T7 DNA polymerase are numbered as numbered by Dunn and Studier, *J. Mol. Biol.* 166:477-535 (1983).

Amino acid residues of Taq DNA polymerase are as numbered in U.S. 5,079,352.

Amino acid residues of the Klenow fragment of *E. coli* are as numbered by Joyce, C. M. *et al.*, *J. Biol. Chem.* 257:1958-1964 (1982).

Amino acid residues of *Thermatoga neapolitana* (Tne) are numbered as in U.S.S.N. 08/370,170, filed January 9, 1995, which is specifically incorporated herein by reference.

Amino acid residues of *Thermatoga maritima* (Tma) DNA polymerase are numbered as in U.S. Patent No. 5,374,553.

In addition to the DNA polymerases mentioned above, it is also possible to prepare the following mutant DNA polymerases:

<u>Enzyme or source</u>	<u>Mutation position</u>
<i>E. coli</i> DNA polymerase I	762
<i>Streptococcus pneumoniae</i>	711
<i>Thermus aquaticus</i>	667
<i>Thermus flavus</i>	666
<i>Thermus thermophilus</i>	669
<i>Deinococcus radiodurans</i>	747
<i>Bacillus caldotenax</i>	711
<i>E. coli</i> bacteriophage T5	570
mycobacteriophage L5	438
<i>E. coli</i> bacteriophage SP01	692
<i>E. coli</i> bacteriophage SP02	447
<i>Thermatoga neapolitana</i>	67 [Figure 4]
<i>Thermatoga maritima</i>	730

The change in amino acid at the mutation positions above is from phenylalanine to tyrosine except for bacteriophage SP02, where the change is from

leucine to tyrosine. Coordinates are as used by Polesky, A.H. *et al.*, *J. Biol. Chem.* 265:14579-14591 (1990) and Astatke M *et al.*, *J. Biol. Chem.* 270:1945-1954 (1995).

The following terms are defined in order to provide a clear and consistent understanding of their use in the specification and the claims. Other terms are well known to the art so that they need not be defined herein.

"Structural gene" is a DNA sequence that is transcribed into messenger RNA and is then translated into a sequence of amino acid residues characteristic of a specific polypeptide.

"Soluble" refers to the physical state of a protein upon expression in a host cell, i.e., the protein has the ability to form a solution *in vivo*. As used herein, a protein is "soluble" if the majority (greater than 50%) of the protein produced in the cell is in solution and is not in the form of insoluble inclusion bodies.

"Nucleotide" is a monomeric unit of DNA or RNA consisting of a sugar moiety, a phosphate, and a nitrogenous heterocyclic base. The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of the pentose). The combination of a base and a sugar is called a nucleoside. Each nucleotide is characterized by its base. The four DNA bases are adenine (A), guanine (G), cytosine (C), and thymine (T). The four RNA bases are A, G, C and uracil (U).

"Processive" is a term of art referring to an enzyme's property of acting to synthesize or hydrolyze a polymer without dissociating from the particular polymer molecule. A processive DNA polymerase molecule can add hundreds of nucleotides to a specific nucleic acid molecule before it may dissociate and start to extend another DNA molecule. Conversely, a non-processive polymerase will add as little as a single nucleotide to a primer before dissociating from it and binding to another molecule to be extended. For the purposes of the present invention, processive refers to enzymes that add, on the average, at least 100, and preferably, about 200 or more, nucleotides before dissociation.

"Thioredoxin" is an enzyme well known to the art that is involved in oxidation and reduction reactions. It is also required as a subunit for T7 DNA

polymerase activity. "Thioredoxin-independent" refers to the ability of and polymerase to be processive in the absence of thioredoxin.

"Promoter" is a term of art referring to sequences necessary for transcription. It does not include ribosome binding sites and other sequences primarily involved in translation.

"Gene" is a DNA sequence that contains information necessary to express a polypeptide or protein. A gene may include homologous or heterologous control elements such as promoters, enhancers, and ribosome binding sites.

"Heterologous" refers herein to two molecules having different origins; i.e. not, in nature, being genetically or physically linked to each other. "Heterologous" also describes molecules that while physically or genetically linked together in nature, are linked together in a substantially different way than is found in nature.

"Homology", as used herein, refers to the comparison of two different nucleic acid sequences. For the present purposes, assessment of homology is as a percentage of identical bases, not including gaps introduced into the sequence to achieve good alignment. Per cent homology may be estimated by nucleic acid hybridization techniques, as is well understood in the art as well as by determining and comparing the exact base order of the two sequences.

"Mutation" is any change that alters the DNA or amino acid sequence. As used herein, a mutated sequence may have single or multiple changes that alter the nucleotide sequence of the DNA or the amino acid sequence of the protein. Alterations of the DNA or amino acid sequence include deletions (loss of one or more nucleotides or amino acids in the sequence), substitutions (substituting a different nucleotide or amino acid for the original nucleotide or amino acid along the sequence) and additions (addition of new nucleotides or amino acids in the original sequence).

"Purifying" refers herein to increasing the specific activity of an enzyme over the level produced in a culture in terms of units of activity per weight of

protein. This term does not imply that a protein is purified to homogeneity. Purification schemes for DNA polymerases are known to the art.

"Expression" is the process by which a polypeptide is produced from a structural gene. It includes transcription of the gene into messenger RNA (mRNA) and the translation of such mRNA into polypeptide(s).

"Substantially pure" means that the desired purified molecule, *e.g.*, enzyme or polypeptide, is essentially free from contaminating cellular components which are associated with the desired enzyme or polypeptide in nature. Contaminating cellular components may include, but are not limited to, phosphatases, exonucleases, endonucleases or other amino acid sequences normally associated with the desired enzyme or polypeptide.

"Origin of replication" refers to a DNA sequence from which DNA replication is begun, thereby allowing the DNA molecules which contain said origin to be maintained in a host, *i.e.*, replicate autonomously in a host cell.

"Host" is any prokaryotic or eukaryotic microorganism that is the recipient of a DNA molecule. The DNA molecule may contain, but is not limited to, a structural gene, expression control elements, *e.g.* a promoter and/or an origin of replication.

"3'-to-5' exonuclease activity" is an enzymatic activity well known to the art. This activity is often associated with DNA polymerases, and is thought to be involved in a DNA replication "editing" or correction mechanism.

"5' to 3' exonuclease activity" is also an enzymatic activity well known in the art. This activity is often associated with DNA polymerases, such as *E. coli* PolI and PolIII.

A "DNA polymerase substantially reduced in 3'-to-5' exonuclease activity" is defined herein as either (1) a mutated DNA polymerase that has about or less than 10%, or preferably about or less than 1%, of the 3'-to-5' exonuclease activity of the corresponding unmutated, wild-type enzyme, or (2) a DNA polymerase having a 3'-to-5' exonuclease specific activity which is less than about 1 unit/mg protein, or preferably about or less than 0.1 units/mg protein. A unit of

activity of 3'-to-5' exonuclease is defined as the amount of activity that solubilizes 10 nmoles of substrate ends in 60 min. at 37°C, assayed as described in the "BRL 1989 Catalogue & Reference Guide", page 5, with *Hha*I fragments of *lambda* DNA 3'-end labeled with <sup>3</sup>[H]dTTP by terminal deoxynucleotidyl transferase (TdT). Protein is measured by the method of Bradford, *Anal. Biochem.* 72:248 (1976). As a means of comparison, natural, wild-type T5-DNAP or T5-DNAP encoded by pTTQ19-T5-2 has a specific activity of about 10 units/mg protein while the DNA polymerase encoded by pTTQ19-T5-2(Exo<sup>-</sup>) (U.S. 5,270,179) has a specific activity of about 0.0001 units/mg protein, or 0.001% of the specific activity of the unmodified enzyme, a 10<sup>5</sup>-fold reduction.

A "DNA polymerase substantially reduced in 5'-to-3' exonuclease activity" is defined herein as either (1) a mutated DNA polymerase that has about or less than 10%, or preferably about or less than 1%, of the 5'-to-3' exonuclease activity of the corresponding unmutated, wild-type enzyme, or (2) a DNA polymerase having 5'-to-3' exonuclease specific activity which is less than about 1 unit mg protein, or preferably about or less than 0.1 units/mg protein.

Both of these activities, 3'-to-5' exonuclease activity and 5'-to-3' exonuclease activity, can be observed on sequencing gels. Active 5'-to-3' exonuclease activity will produce nonspecific ladders in a sequencing gel by removing nucleotides from growing primers. 3'-to-5' exonuclease activity can be measured by following the degradation of radiolabeled primers in a sequencing gel. Thus, the relative amounts of these activities, e.g. by comparing wild-type and mutant polymerases, can be determined from these characteristics of the sequencing gel.

As used herein, "amplification" refers to any *in vitro* method for increasing the number of copies of a nucleotide sequence with the use of a DNA polymerase. Nucleic acid amplification results in the incorporation of nucleotides into a DNA molecule or primer, thereby forming a new DNA molecule complementary to a DNA template. The formed DNA molecule and its template can be used as templates to synthesize additional DNA molecules. As used herein, one

amplification reaction may consist of many rounds of DNA replication. DNA amplification reactions include, for example, polymerase chain reaction (PCR). One PCR reaction may consist of 30-100 "cycles" of denaturation and synthesis of a DNA molecule.

As used herein, "thermostable" refers to a DNA polymerase which is resistant to inactivation by heat. DNA polymerases synthesize the formation of a DNA molecule complementary to a single-stranded DNA template by extending a primer in the 5'-to-3' direction. This activity for mesophilic DNA polymerases may be inactivated by heat treatment. For example, T5 DNA polymerase activity is totally inactivated by exposing the enzyme to a temperature of 90°C for 30 seconds. As used herein, a thermostable DNA polymerase activity is more resistant to heat inactivation than a mesophilic DNA polymerase. However, a thermostable DNA polymerase does not mean to refer to an enzyme which is totally resistant to heat inactivation, and thus heat treatment may reduce the DNA polymerase activity to some extent. A thermostable DNA polymerase typically will also have a higher optimum temperature than mesophilic DNA polymerases.

The present invention is directed to a recombinant DNA molecule having a mutated DNA sequence encoding a protein which has DNA polymerase activity and which incorporates dideoxynucleotides about as well as deoxynucleotides. The mutant DNA molecule of the invention may also contain expression control elements, *e.g.* a promoter and/or an origin of replication. In this combination, a promoter and the structural gene are positioned and orientated with respect to each other such that the structural gene may be expressed in a host cell under the control of the promoter. The origin of replication is capable of maintaining the promoter/structural gene/origin of replication combination in a host cell. Preferably, the promoter and the origin of replication are functional in the same host cell, such as an *E. coli* host cell. The DNA molecule is preferably a transformed host cell, exemplified herein by an *E. coli* host cell (in particular, *E. coli* DH10B), but may also exist *in vitro*. The promoter may be any constitutive or inducible promoter. Examples of constitutive promoters that may be used in

00958421 "042600"

8-12/86



the practice of the invention include ribosomal protein promoter, RPSL, and the ampicillin resistance gene promoter. Examples of inducible promoters include the *lambda* P<sub>L</sub> promoter, *tac* promoter, and *lac* promoter. The expressed protein of the invention may have a processive 3'-to-5' DNA exonuclease activity or may have substantially reduced 3'-to-5' exonuclease activity. The expressed protein of the invention may also have a 5'-to-3' DNA exonuclease activity or may have substantially reduced 5'-to-3' DNA exonuclease activity. The expressed protein of this invention may also have both substantially reduced processive 3'-to-5' DNA exonuclease activity and substantially reduced 5'-to-3' DNA exonuclease activity. Preferably, the structural gene is expressed under the control of a heterologous promoter. In addition, the structural gene may be expressed under the control of a heterologous ribosome binding site, although the native DNA polymerase ribosomal binding site may also be used.

The present invention pertains both to the mutant DNA polymerase and to its functional derivatives. The term "functional derivative" is intended to include the "fragments," "variants," "analogues," and "chemical derivatives" of a molecule. A "fragment" of a molecule such as a DNA polymerase, is meant to refer to any polypeptide subset of the molecule. A "variant" of a molecule such as a DNA polymerase is meant to refer to a molecule substantially similar in structure and function to either the entire molecule, or to a fragment thereof. A molecule is said to be "substantially similar" to another molecule if both molecules have substantially similar structures or if both molecules possess a similar biological activity. Thus, provided that two molecules possess a similar activity, they are considered variants as that term is used herein even if the structure of one of the molecules is not found in the other, or if the sequence of amino acid residues is not identical. An "analogue" of a molecule such as a DNA polymerase is meant to refer to a molecule substantially similar in function to either the entire molecule or to a fragment thereof. As used herein, a molecule is said to be a "chemical derivative" of another molecule when it contains additional chemical

moieties not normally a part of the molecule. Such moieties may improve the molecule's solubility, absorption, biological half life, etc.

5 The present invention also relates to a method for the production of a protein having a mutant DNA polymerase activity as described herein by the steps of culturing a cell containing a mutant DNA molecule of the invention under conditions where the DNA is expressed, followed by purifying the protein expressed during the culturing step. In this method, the recombinant DNA molecule encodes the protein, and also includes a promoter and an origin of replication. (The promoter and the structural gene are in such position and orientation with respect to each other that the promoter may regulate the expression of the gene in the cell). The origin of replication may be heterologous to the structural gene and capable of maintaining the structural gene/promoter/origin of replication combination in the host cell. Preferably, the mutant DNA polymerase gene is expressed and maintained in an *E. coli* host cell. 10 The promoter may be heterologous to the structural gene and may be inducible, e.g. a *lambda* P<sub>L</sub> promoter, a *tac* promoter, or a *lac* promoter. Preferably, the structural gene is under control of a heterologous promoter. The structural gene of the invention may be under control of a heterologous ribosome binding site. The protein may have a processive 3'-to-5' DNA exonuclease activity or may have substantially reduced 3'-to-5' exonuclease activity. The protein may also have 5'-to-3' exonuclease activity or may have substantially reduced 5'-to-3' exonuclease activity. The protein may have both substantially reduced 3'-to-5' exonuclease activity and substantially reduce 5'-to-3' exonuclease activity. 20

25 Although specific plasmids, vectors, promoters and host cells are disclosed and used in the Example section, other promoters, vectors, and host cells, both prokaryotic and eukaryotic, are well known in the art and in keeping with the specification, may be used to practice the invention. Eukaryotic cells include yeast, CHO, and BHK. Prokaryotic cells include *E. coli*, *Samonella*, *Baccillus* and *Streptomyces*. Specific molecules exemplified herein include pTTQ-Taq, pSport5-3, pUC-TneFY, pTrcTne35FY, pTTQTne535FY, pTTQTne5FY, and 30

pTrcTneFY, and functional derivatives thereof. A functional derivative of a DNA molecule is derived from the original DNA molecule but still may express the desired mutant DNA polymerase structural gene in a host or *in vitro* according to the present invention.

5 The present invention further relates to a mutant DNA polymerases produced by the method of the present invention, having substantially reduced exonuclease activities. Standard protein purification techniques well known in the art may be used to purify the polymerase proteins of the present invention. Preferably, the exonuclease activity is less than about 1 unit/mg protein. More preferably, the exonuclease activity is less than about 0.1 units/mg protein. Even more preferably, the exonuclease activity is less than about 0.003 units/mg protein. Most preferably, the exonuclease activity is less than about 0.0001 units/mg protein.

10 The amino acid sequences of the DNA polymerases were compared with other known DNA polymerases, such as *E coli* DNA polymerase I, Taq DNA polymerase, T5 DNA polymerase, and T7 DNA polymerase to localize the regions of 3'-to-5' exonuclease activity as well as the polymerase and dNTP binding domains. Based on this comparison of the amino acid sequences of various DNA polymerases (Blanco *et al.*, Gene 112:139-144 (1992); Braithwaite and Ito, Nucleic Acids Res. 21:787-802 (1993)), a 3'-to-5' exonuclease domain was localized as follows:

15

20

25

*			
Tne	317	PSFALDLETSS	327 <sup>1</sup>
Pol I	350	PYFAFDTETDS	360
T5	133	GPVAFDSETSA	143
T7	1	-MIVSDIEANA	10

Mutations, such as insertions, deletions, and substitutions, within this domain can result in substantially reduced 3'-to-5' exonuclease activity. By way

---

<sup>1</sup> Numbering is as reported in U.S.S.N. 08/370,190, filed January 9, 1995.

of example, Asp<sup>322</sup> (Tne), Asp<sup>355</sup> (Pol I), Asp<sup>138</sup> (T5), and Asp<sup>5</sup> (T7), noted by an \*, can be converted to Ala<sup>322</sup>, Ala<sup>355</sup>, Ala<sup>138</sup>, and Ala<sup>5</sup>, respectively, to obtain mutants with substantially reduced 3'-to-5' exonuclease activity.

5 The mutant DNA polymerase of the invention may have 5'-to-3' exonuclease activity or may have substantially reduced 5'-to-3' exonuclease activity. In most of the known polymerases, the 5'-to-3' exonuclease domain is present at the N-terminal region of this polymerase. Ollis, DL *et al.*, *Nature* 313:762-766 (1985); Freemont *et al.*, *Protein* 1:66-73 (1986); Joyce, C.M., *Curr. Opin. Struct. Biol.* 1:123-129 (1991). There are some conserved amino acids that have been implicated as responsible for 5'-to-3' exonuclease activity. Gutman and Minton, *Nucl. Acids Res.* 21:4406-4407 (1993). <sup>In *E. coli* Pol I the</sup> ~~These~~ amino acids include Tyr<sup>77</sup>, Gly<sup>103</sup>, Gly<sup>184</sup> and Gly<sup>192</sup>. The 5'-to-3' exonuclease domain is dispensable. The best known example is the Klenow fragment of *E coli* polymerase I. The Klenow fragment is a natural proteolytic fragment devoid of 5'-to-3' exonuclease activity. Joyce, C.M., et al., *J. Biol. Chem.* 257:1958-64 (1990). For example, the 219 N-terminal amino acid residues of the Tne DNA polymerase can be deleted to result in a mutant with substantially diminished 5'-to-3' exonuclease activity.

10 The mutant DNA polymerases of this invention may be used in cloning and *in vitro* gene expression experiments to produce heterologous polypeptides from the cloned genes. The mutant-DNA polymerases of this invention may also be used for DNA sequencing, DNA labeling, and amplification reactions.

15 As is well known, sequencing reactions, such as dideoxy DNA sequencing in cycle DNA sequencing of plasmid DNA, require the use of DNA polymerases. Dideoxy-mediated sequencing involves the use of a chain-termination technique which uses a specific polymer for extension of DNA polymerase, a base-specific chain terminator, and the use of polyacrylamide gels to separate the newly synthesized chain-terminated DNA molecules by size so that at least a part of the nucleotide sequence of the original DNA molecule can be determined. Specifically, a DNA molecule is sequenced by using four separate DNA sequencing reactions, each of which contains different base-specific terminators.

20

25

30

D.C. 12/8/95

For example, the first reaction will contain a G-specific terminator, the second reaction will contain a T-specific terminator, the third reaction will contain an A-specific terminator, and a fourth reaction may contain a C-specific terminator. Preferred terminator nucleotides include dideoxyribonucleoside triphosphates (ddNTPs) such as ddATP, ddTTP, ddGTP, and ddCTP. Analogues of dideoxyribonucleoside triphosphates may also be used and are well known in the art.

When sequencing a DNA molecule, ddNTPs lack a hydroxyl residue at the 3' of the deoxyribonucleoside, and thus, although they can be incorporated by DNA polymerases into the growing DNA chain, the absence of the 3' hydroxyl residue prevents formation of a phospho-diester bond, resulting in termination of extension of the DNA molecule. Thus, when a small amount of one ddNTP is included in a sequencing reaction mixture, there is a competition between extension of the chain and base-specific termination, resulting in a population of synthesized DNA molecules which are shorter in length than the DNA template to be sequenced. By using four different ddNTPs and four separate enzymatic reactions, populations of the synthesized DNA molecules can be separated by size so that at least a part of the nucleotide sequence of the original DNA molecule can be determined. DNA sequencing by dideoxy nucleotides is well known, and is described by Sambrook *et al.* in: *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989). As will be readily recognized, the DNA polymerases of the present invention may be used in such sequencing reactions.

As is well known, detectably labeled nucleotides are typically included in sequencing reactions. Any number of labeled nucleotides can be used in sequencing or labeling reactions, including, but not limited to, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels, and enzyme labels. It has been unexpectedly discovered that the Tne DNA polymerase of the present invention may be particularly useful for incorporating  $\alpha$ S nucleotides during sequencing or labeling reactions. For example,  $\alpha^{35}\text{[S]dATP}$ ,

a commonly-used detectably-labeled nucleotide in sequencing reactions, is incorporated three times more efficiently with the Tne polymerase of the present invention than with Taq DNA polymerase. Thus, the enzymes of the present invention are suited for sequencing or labeling DNA molecules with  $\alpha^{35}$ dNTPs. Particularly suited is Tne DNA polymerase or mutants thereof.

Polymerase chain reaction (PCR), a well-known DNA amplification technique, is a process by which DNA polymerase and deoxyribonucleoside triphosphates are used to amplify a target DNA template. In such PCR reactions, two primers, one complementary to the 3' termini (or near the 3' termini) of the first strand of the DNA molecule to be amplified, and a second primer complementary to the 3' termini (or near the 3' termini) of the second strand of the DNA molecule to be amplified, are hybridized to their respective DNA molecules. After hybridization, DNA polymerase, in the presence of deoxyribonucleoside triphosphates, allows the synthesis of a third DNA molecule complementary to the first strand, and a fourth DNA molecule complementary to the second strand of the DNA molecule to be amplified. This synthesis results in two double-stranded DNA molecules. Such double-stranded DNA molecules may then be used as DNA templates for synthesis of additional DNA molecules by providing a DNA polymerase, primers, and deoxyribonucleoside triphosphates. As is well known, the additional synthesis is carried out by "cycling" the original reaction (with excess primers and deoxyribonucleoside triphosphates), allowing multiple denaturing and synthesis steps. Typically, denaturing of double-stranded DNA molecule to form single-stranded DNA templates is accomplished by high temperatures. For example, the Tne DNA polymerase of the present invention is a heat-stable DNA polymerase and thus will survive such thermocycling during DNA amplification reactions. Thus, the Tne DNA polymerase is suited for PCR reactions, particularly where high temperatures are used to denature the DNA molecules during amplification. In addition to Taq, wild type and mutant *Thermus flavus*, *Thermus thermophilus*, and *Thermus aquaticus* DNA polymerases are useful for PCR.

The DNA polymerases of the invention are ideally suited for the preparation of a kit. Kits comprising the DNA polymerase may be used to detectably label DNA molecules, for DNA sequencing, or for DNA amplification by well-known techniques. Such kits may comprise a carrying means being compartmentalized to receive, in close confinement, one or more container means such as vials, test tubes, and the like. Each of such container means comprises components or a mixture of components needed to perform DNA sequencing, DNA labeling, or DNA amplification.

A kit for sequencing DNA may comprise a number of container means. A first container means may, for example, comprise a substantially purified DNA polymerase of the invention. A second container means may comprise one or a number of types of nucleotides needed to synthesize a DNA molecule complementary to a DNA template. A third container means may comprise one or a number of different types of ddNTPs. In addition to the above container means, additional container means may be included in the kit comprising one or a number of DNA primers.

A kit used for amplifying DNA will comprise, for example, a first container means comprising a substantially pure DNA polymerase and one or a number of additional container means which comprise a single type of nucleotide or mixtures of nucleotides. Various primers may or may not be included in a kit for amplifying DNA.

When desired, the kit of the present invention may also include container means which comprise detectably labeled nucleotides which may be used during the synthesis or sequencing of a DNA molecule. One or a number of labels may be used to detect such nucleotides. Illustrative labels include, but are not limited to, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels, and enzyme labels.

Having now generally described this invention, the same will be better understood by reference to specific examples, which are included herein for

purposes of illustration, and are not intended to be limiting unless otherwise specified.

### *Examples*

The overall cloning strategy used in the Examples may be more easily understood by reference to the Figures.

#### *Example 1: Preparation of Non-Discriminating Mutant DNA Polymerases*

As models, T5, Tne, and Taq DNA polymerases were used. The polymerase active site, including the dNTP binding domain, is usually present in the C-terminal region of the polymerase (Ollis, D.L., *et al.*, *Nature* 313:763-766 (1985); Freemont, P.S., *et al.*, *Proteins* 1: 66-73 (1986).) Our partial sequence of the Tne polymerase gene suggests that the amino acids that presumably contact and interact with the dNTPs are present within the 694 bases starting of the internal *Bam*HI site, based on the homology with the prototype polymerase *E. Coli* PolII (Poleskey A.H., *et al.*, *J. Biol. Chem.* 265:14579-14591 (1990). The corresponding amino acids in other polymerases are present in the O helix.

Initially, it was attempted to replace amino acids 544 to 729 (coordinates from Leavitt and Ito, *Proc. Natl. Acad. Sci USA* 86:4465-4469 (1989)) of T5 DNA polymerase with amino acids 500 to 675 (coordinates from Dunn and Studier, *J. Mol. Biol.* 166: 477-535 (1983)) of T7 DNA polymerase. This region encompasses entire O-helix plus additional amino acids on either side of the helix. The extra amino acids were chosen for convenient restriction sites *Dra*III and *Ssp*I present in T5 DNA polymerase. The corresponding region of T7 was generated by PCR using the oligos:

[SEQ. ID. No. 1]:

5'- CAGGATCCACATGGTGCTTAACGGCGACATCCACACTAAG and



[SEQ. ID. No. 2]:

GTAACTTCTTGTGCGGTCTCAATGAC.

The hybrid plasmid containing the active sites of T7 DNA polymerase was constructed by replacing the T5 active sites with the PCR product. However, the construct did not produce any active protein, perhaps because the structure of the altered protein was unstable in *E. coli*. Therefore, it was reasoned that it may be possible to change specific amino acids of T5 DNA polymerase and Taq DNA polymerase in the O-helix (based on the sequence comparison with the T7 DNA polymerase) to produce an active hybrid polymerase. This small change should not alter significantly the structure of the mutant polymerase.

The amino acid sequence in the O-helix of T7, T5, Tne, Taq, and the Klenow fragment are as follows:

Tma	725	GKMVNFSIIYG	735 [SEQ ID No. 17]
T5	565	AKAITFGILYG	675 [SEQ ID No. 3]
T7	521	AKTFIYGFLYG	531 [SEQ ID No. 4]
Taq	662	AKTINFGVLYG	672 [SEQ ID No. 5]
Klenow frag.	757	AKAINFGLIYG	767 [SEQ ID No. 6]
Tne	62	GKMVNFSIIYG	72 [SEQ ID No. 12]

The sequence of the Klenow fragment is disclosed by Polesky, A.H. *et al.*, *J. Biol. Chem.* 265:14579-14591 (1990), and the sequence of the C-terminal portion of the Tne polymerase gene is shown in Figure 4.

T7 DNA polymerase has a sequence stretch Thr-Phe-Ile-Tyr [SEQ ID No. 7] in the O-helix. The corresponding sequence in T5, Taq, and Tne DNA polymerase are Ala-Ile-Thr-Phe [SEQ ID No. 8]; Thr-Ile-Asn-Phe [SEQ ID No. 9]; and Met-Val-Asn-Phe [SEQ ID No. 13], respectively. These amino acid are bordered by known conserved dNTP binding amino acids Lys (K) and Tyr (Y). Therefore, it was tested whether changing these amino acids of T5 and Taq DNA polymerases to Thr-Phe-Ile-Tyr [SEQ ID No. 7] would make the polymerases as non-discriminating as T7 DNA polymerase. One of the main differences in this

region of T7 DNA polymerase is that it contains a tyrosine residue with an hydroxyl group in place of phenylalanine in the case of Klenow fragment, T5 and Taq DNA polymerases. An oligo T CAG GCT GCT AAA ACA TTC ATC TAC GGT ATA CTG TAT GGT TCT GG [SEQ ID No. 10] was generated to change Ala-Ile-Thr-Phe [SEQ ID No. 8] of T5 DNA polymerase to Thr-Phe-Ile-Tyr [SEQ ID No. 7] by site directed mutagenesis. The oligo was also designed to create an *AccI* site to detect the mutant clone in the process. The mutagenesis was done using BioRad Mutagene Kit (BioRad, California) according to the protocol described by the manufacturer.

Protocol for mutagenesis: pSport T5-E (Fig. 1) was digested with *Clal*I and *EcoRI* to generate a 1.9 kb fragment of T5 DNA polymerase. The fragment was cloned onto M13mp18 (LTI, Gaithersburg, MD) at the *AccI* and *EcoRI* sites. The recombinant clone was selected in DH5 $\alpha$ F'IQ (LTI, Gaithersburg, MD). Single stranded uracilated DNA was isolated from CJ236 (Biorad, California) and used for site-directed mutagenesis using the Biorad Mutagene kit. Following mutagenesis, 6 clones were tested for the presence of an additional *AccI* site included in the mutagenic oligo [SEQ ID No. 7]. Five of the six clones produced about a 1 kb fragment, an indication that these clones contain expected mutations. One of the clones was used to replace the wild type fragment in pSport T5-E. First, a 1.0 kb *DraIII-EcoRI* fragment of wild type T5 DNA polymerase of pUC#3-Exo (which contains the *BamHI-EcoRI* fragment of T5 DNA polymerase gene from pSportT5-E in pUC19) replaced with the *DraIII-EcoRI* fragment of the mutant M13 RF DNA. This fragment contains the mutations. Second, *EcoRI*-6 fragment T5 Phage DNA which contains the residual COOH-end of the T5 polymerase gene (U.S. Patents 5,270,179 and 5,047,342) was cloned in pUCT5 mutant in order to reconstruct the entire T5 DNA polymerase gene. Finally, the entire T5 polymerase gene containing the mutations (A1TF to TFIY) was cloned in pSport (LTI, Gaithersburg, MD) at the *BamHI* site as *BamHI-BglII* as described before (U.S. Patent Nos. 5,270,170 and 5,047,342). An active hybrid (T5/T7) polymerase was obtained from the recombinant clone.

Similarly, an oligo GTA GAG GAC CCC GTA ATT AAT GGT CTT GGC CGC [SEQ ID No. 11] was designed to change the phenylalanine residue (amino acid 667) to a tyrosine of Taq DNA polymerase. An *AseI* site was also created for initial screening of the mutant clones.

5 Since thermostable Taq DNA polymerase cloned and expressed in *E. coli* can be purified very easily, the mutant Taq DNA polymerase was characterized with respect to its DNA polymerase activity and its ability to produce sequencing ladder in the presence of varying amount of dideoxynucleotides. A 2.5 kb portion of Taq DNA polymerase (Fig. 2) was cloned as a *HindIII-XbaI* fragment in M13mp19 (LTI, Gaithersburg, MD). The kinased oligo was used for mutagenesis by the procedure as described above. Following mutagenesis the mutant fragment was cloned in the expression vector as follows.

10 The DNA fragment from the mutant phage DNA was obtained by digesting the DNA with *NgoAIV* and *XbaI*. The 1.6 kb *NgoAIV-XbaI* fragment of pTTQ-Taq was replaced with the 1.6 kb *NgoAIV-XbaI* fragment containing the mutation (F667Y). The mutant clone produced active polymerase.

15 Upon testing in the DNA sequencing reaction optimized for wild type Taq DNA polymerase, it was found that the mutant Taq DNA polymerase is unable to produce satisfactory ladder and the DNA sequencing (synthesis) is terminated prematurely. With the mutant Taq DNA polymerase, only nine bases of pUC18 DNA sequence were able to be read. The wild type Taq DNA polymerase produced expected sequencing ladder under identical conditions (we were able to read up to 300-400 bases of pUC18 DNA sequence). This is an indication that the mutant polymerase is incorporating dideoxynucleotides very efficiently and the DNA sequencing reaction is terminating prematurely. By decreasing the concentration of dideoxynucleoside triphosphates, it was possible to generate a sequencing ladder. (When dNTPs concentrations were held constant to 20  $\mu$ M, and the ddNTP concentrations were reduced 100-fold, >400 bases of sequence in the G-lane (0.4  $\mu$ M) were read. The other ddNTPs needed to be reduced even further as their initial concentrations were 5-10 fold higher.) The mutant Taq

20

25

30

DNA polymerase needed 100-fold less dideoxynucleotides compared to the wild type Taq DNA polymerase to generate DNA sequencing ladder. This suggests that the mutant Taq DNA polymerase became nondiscriminatory upon modification of phenylalanine 667 to tyrosine 667. In addition, the mutant Taq DNA polymerase produced almost uniform band intensity compared to the wild-type in the sequencing ladder in the presence of chain terminating dideoxynucleotides. The result suggests that uneven band intensity in the sequencing ladder was at least in part due to discriminatory activity towards nonnatural nucleotides.

An attempt was made to generate a similar mutant of Tne DNA polymerase. It was anticipated that the mutant Tne DNA polymerase will be better in DNA sequencing reactions because Tne DNA polymerase inherently incorporates [ $\gamma$ - $^{35}$ S] dNTPs 3-to-5 fold better than the Taq polymerase, a property highly desirable in DNA sequencing. In order to change the Phe<sup>67</sup> to a Tyr<sup>67</sup> (Fig. 4; SEQ ID No. 14) site-directed mutagenesis was performed using the oligonucleotide.

GTA TAT TAT AGA GTA GTT AAC CAT CTT TCCA. [SEQ ID No. 15]

In this oligonucleotide a *HpaI* restriction site was introduced to facilitate screening of the mutants. To make a mutant Tne DNA polymerase, a 2kb *SphI* fragment of pSport -Tne (Figure 3) was cloned into M13mp19 (LTI, Gaithersburg, MD). The recombinant was selected in *E. coli* DH5aF'IQ (LTI, Gaithersburg, MD). One of the clones with a proper insert was used to isolate uracilated single-stranded DNA by infecting *E. coli* CJ236 (Biorad, California), with a phage particle obtained from *E. coli* DH5aF'IQ. A single-stranded uracilated DNA was used for site-directed mutagenesis using the protocol described in the BioRad manual, see *supra*, except T7 DNA polymerase was used instead of T4 DNA polymerase. The resulting mutants were screened for the presence of the *HpaI* site. Mutants with the desired *HpaI* site were used for further study.

DNA containing the Phe<sup>67</sup>→Tyr<sup>67</sup> mutations were incorporated into pUC-Tne by replacing the wild type *SphI*-*HindIII* fragment with the mutant fragment obtained from the mutant phage DNA from the site-directed mutagenesis. The structure of the desired clone, pUC-Tne FY, was confirmed by the presence of the unique *HpaI* site. (Fig. 5A) The entire mutant polymerase gene was subcloned into pTrc99. The plasmid, pUC-TneFY, was digested with *SstI* and *HindIII* and the entire mutant polymerase gene (2.6kb) was purified and cloned with *SstI* and *HindIII* digested pTrc99 expression vector (Pharmacia, Sweden). The clones were selected in DH10B (LTI, Gaithersburg, MD). The desired plasmid was designated pTrcTneFY (Fig. 5B). The clone produced active heat stable polymerase.

The purification of the mutant Tne polymerase was done essentially as described in U.S. Patent Application, Serial No. 08/370,190, filed January 9, 1995, incorporated by reference herein, with minor modifications. Five to 10 grams of cells expressing the cloned mutant Tne DNA polymerase were lysed by sonication with a Heat Systems Ultrasonic Inc. Model 375 sonicator in a sonication buffer consisting of 50 mM Tris-HCl, pH 7.4, 8% glycerol, 5mM 2-mercaptoethanol, 10mM NaCl, 1mM EDTA, 0.5 mM PMSF. The sonicated sample was heated at 75°C for 15 min. Following heat treatment, 200 mM NaCl and 0.4% PEI was added to remove nucleic acids. The extract was centrifuged for clarification. Ammonium sulfate was added to 48%, the pellet was resuspended in a column buffer consisting of 25 mM Tris-HCl, pH 7.4, 8% glycerol, 0.5% EDTA, 5 mM 2-mercaptoethanol, 10 mM KCl, and loaded on a Heparin column. The column was washed with 10 column volumes of a buffer gradient from 10 mM to 1 M KCl. Fractions containing polymerase activity were pooled and dialyzed in column buffer as above except the pH is 7.8. The dialyzed pooled fractions were loaded onto a MonoQ column. The column was washed and eluted as described above. The active fractions are pooled and a unit assay was done. The reaction contained 25 mM TAPS, pH 9.3, 2 mM MgCl<sub>2</sub>, 50 mM KCl, 1 mM DTT, 0.2 mM dNTPs, 500 µg/ml DNase I treated salmon sperm

DNA, 21 mCi/ml [ $\alpha$ P<sup>32</sup>] dCTP and various amount of polymerase in a final volume of 50  $\mu$ l. After 10 min. at 70°C, 10  $\mu$ l of 0.5 M EDTA was added to the tube. TCA precipitable counts were measured in GF/C filters using 40  $\mu$ l of the reaction.

5           Upon testing in the DNA sequencing reaction, the TneFY mutant polymerase gave only a 9 base sequencing ladder when the Taq cycle sequencing reaction conditions were used (LTI). Diluting the dideoxynucleotides by a factor of 100 extended the ladder to about 200 bases. The F $\rightarrow$ Y mutation in the TneFY polymerase, therefore, allowed dideoxynucleotides to be incorporated at a much higher frequency than for wild-type polymerase. Taken together, it can be concluded that T5, Taq, Tne, Tma and other DNA polymerases can be made nondiscriminatory towards dideoxynucleotide and perhaps other nonnatural nucleotides by simple modification of a specific phenylalanine residue to a tyrosine residue. These DNA polymerases are useful in DNA sequencing and other molecular biological applications.

***Example 2: Preparation of Non-Discriminating Mutant DNA Polymerase Substantially Reduced in 3'-to-5' Exonuclease Activity***

20           To make the 3'-to-5' exonuclease mutants, an oligonucleotide, GA CGT TTC AAG CGC TAG GGC AAA AGA [SEQ ID No. 16] was used to convert the Asp<sup>322</sup> to Ala<sup>322</sup>. An Eco47III site was created to facilitate screening of the mutant following mutagenesis. The mutagenesis was performed using a protocol as described in the Biorad manual except T7 DNA polymerase was used instead of T4 DNA polymerase. *See supra*. The mutant clones were screened for an Eco47III site that was created in the mutagenic oligonucleotide. One of the  
25           mutants having the created Eco47III site was used for further study.

          To incorporate the 3'-to-5' exonuclease mutation into an expression vector, the mutant phage DNA obtained as described above was digested with *Sph*I and *Hind*III and a 2 kb fragment containing the mutation was isolated. The

fragment was cloned in pUC-Tne to replace the wild-type fragment (Figure 5A). The desired clone, pUC-Tne (3'→5') was confirmed by the presence of a unique Eco47III site. The plasmid digested with *Sst*I and *Hind*III in the entire mutant polymerase gene (2.6 kb) was purified and cloned into *Sst*I and *Hind*III digested pTrc99 expression vector, obtainable from Pharmacia, Sweden. The clones were selected in DH10B (LTI, Gaithersburg, MD). The desired plasmid was designated as pTrcTne35 (Figure 5B). The clone produced active heat stable polymerase. The polymerase was purified as described *supra*, for TneFY in Example 1.

In order to introduce both the 3'-to-5' exonuclease mutation and the Phe<sup>67</sup>→Tyr<sup>67</sup> mutation in the expression vector pTrc99, it was first necessary to reconstitute both mutations in a pUC-Tne clone (Figure 6). Both pUC-Tne (3'-to-5') and pUC-TneFY were digested with *Bam*HI. The digested pUC-Tne (3'→5') was desphosphorylated to avoid recircularization in the following ligation step. Both digested plasmids were run in a 1% agarose gel. The largest *Bam*HI fragment (4.4 kb) was purified from pUC-Tne (3'→5') digested DNA and the small *Bam*HI fragment (0.8 kb) containing the Phe<sup>67</sup>→Tyr<sup>67</sup> mutation was purified and ligated to generate pUC-Tne35FY. The proper orientation and the presence of both mutations were confirmed by *Eco*47III, *Hpa*I, and *Sph*I-*Hind*III restriction digest (Figure 6). Finally, the entire polymerase gene containing both mutations was subcloned as an *Sst*I-*Hind*III fragment in pTrc99 to generate pTrcTne35FY in DH10B. The clone produced active heat stable polymerase. The polymerase was purified as described *supra* to Example 1.

The Tne35FY mutant was used in cycle sequencing reactions using P<sup>32</sup> end-labeled primers. This mutant produced a sequencing ladder and exhibited a similar ability to incorporate dideoxynucleotides as TneFY. In this case the sequence extended to beyond 400 bases and the excess P<sup>32</sup> end-labeled M13/pUC Forward 23-base Sequencing Primer band remained as a 23-base position in the ladder. The persistence of the 23-base primer band confirmed that the 3'-to-5' exonuclease activity had been significantly reduced.

***Example 3: Preparation of Non-Discriminating Mutant DNA Polymerases Exhibiting Substantially Reduced 5'-to-3' Exonuclease Activity***

In order to generate an equivalent mutant devoid of 5'-to-3' exonuclease activity as well as 3'-to-5' exonuclease activity, the presence of a unique *Sph*I site present 680 bases from the *Sst*I site was exploited. pUC-Tne35FY was digested with *Hind*III, filled-in with Klenow fragment to generate a blunt-end, and digested with *Sph*I. The 1.9 kb fragment was cloned into an expression vector pTTQ19 at the *Sph*I-SmaI sites and was introduced into *E. coli* DH10B. (Stark, M.J.R., *Gene* 51:255-267 (1987)). This cloning strategy generated an in-frame polymerase clone with an initiation codon for methionine from the vector. The resulting clone is devoid of 219 amino terminal amino acids of Tne DNA polymerase. This clone is designated as pTTQTne535FY. The clone produced active heat stable polymerase. No exonuclease activity could be detected in the mutant polymerase as evidence by lack primer degradation previously labeled with radioisotope in the sequencing reaction. The mutant polymerase was purified as described *supra*, in Example 1. This particular mutant polymerase is highly suitable for DNA sequencing.

Cycle sequencing reactions using P<sup>32</sup> end -labeled primers were prepared using this mutant. The sequencing reaction produced sequencing ladders. The Tne535FY mutant performed similarly to the Tne35FY mutant except that the signal intensity increased by at least 5 fold. The background was very low in the relative band intensities were extremely even, showing no patterns of sequence-dependent intensity variation.

A 5'-to-3' exonuclease deletion mutant of Tne DNA polymerase containing a Phe<sup>67</sup>→Tyr<sup>67</sup> mutation was also obtained. In order to generate this mutant, the 1.8 kb *Sph*I-*Spe*I fragment (Figure 7) of pTTQTne35FY was replaced with the identical fragment of pUC-TneFY. The clone, pTTQTne5FY, produced active heat stable polymerase. The mutant had modulated, low but detectable, 3'-to-5' exonuclease activity compared to wild-type Tne DNA polymerase as



measured by the rate of degradation of the labeled primer. M13 sequencing primer (LTI, Gaithersburg, MD) was labeled at the 5'-end with [ $\gamma^{32}$ ] ATP and T4 Kinase (LTI, Gaithersburg, MD) as described by the manufacturer. The reaction contained 2.0 units of either wild-type or the mutant Tne DNA polymerase, 0.25 pmol of labeled primer, 20 mM Tricine, pH 8.7, 85 mM potassium acetate, 1.2 mM magnesium acetate, and 8% glycerol. Incubation was carried out at 70°C. At various time points, 10  $\mu$ l aliquots were removed to 5  $\mu$ l cyclesequencing stop solution and resolved in 6% polyacryamide sequencing gel followed by autoradiography. While the wild-type polymerase degraded the primer in 5 to 15 minutes, it took the mutant polymerase more than 60 minutes for the same amount of degradation of the primer. Preliminary results suggest that this particular mutant polymerase is able to amplify more than 12 kb of genomic DNA when used in conjunction with Taq DNA polymerase. Thus, this mutant polymerase will be suitable for large fragment PCR.

Although the foregoing refers to particular preferred embodiments, it will be understood that the present invention is not so limited. It will occur to those of ordinary skill in the art that various modifications may be made to the disclosed embodiments and that such modifications are intended to be within the scope of the present invention, which is defined by the following Claims.

All publications and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

***What Is Claimed Is:***

1. A DNA molecule comprising a coding sequence for a mutant protein, wherein said mutant protein is a mutant DNA polymerase selected from the group consisting of: *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase I, *Streptococcus pneumoniae* polymerase, *Thermus aquaticus* polymerase, *Thermus flavus* polymerase, *Thermus thermophilus* polymerase, *Deinococcus radiodurans* polymerase, *Bacillus caldotenax* polymerase, *E. coli* bacteriophage T5 polymerase, mycobacteriophage L5 polymerase, *Thermatoga maritima* polymerase, and *E. coli* bacteriophage SP01 polymerase, and

wherein said mutant DNA polymerase comprises a substitution of Tyr for Phe at a position in said polymerase corresponding to Phe<sub>570</sub> of wild-type T5 polymerase.

2. The DNA molecule of claim 1, further comprising a promoter, wherein said promoter is in a position and orientation with respect to the coding sequence such that the mutant protein may be expressed in a cell under the control of said promoter.

3. The molecule of claim 2, wherein said coding sequence is heterologous to said promoter.

4. A host cell comprising the DNA molecule of claim 1.

5. The host cell of claim 4, wherein said host cell is *E. coli*.

6. A method for producing a protein, wherein said protein is a mutant DNA polymerase selected from the group consisting of: *E. coli* DNA polymerase

I, Klenow fragment of *E. coli* DNA polymerase I, *Streptococcus pneumoniae* polymerase, *Thermus aquaticus* polymerase, *Thermus flavus* polymerase, *Thermus thermophilus* polymerase, *Deinococcus radiodurans* polymerase, *Bacillus caldotenax* polymerase, *E. coli* bacteriophage T5 polymerase, mycobacteriophage L5 polymerase, *Thermatoga maritima* polymerase, and *E. coli* bacteriophage SP01 polymerase, comprising a substitution of Tyr for Phe at a position in said polymerase corresponding to Phe<sub>570</sub> of wild-type T5 polymerase, said method comprising:

- (a) culturing a host cell comprising the DNA molecule of claim 2, and
- (b) isolating said protein from said host cell.

7. A mutant DNA polymerase selected from the group consisting of a mutant of: *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase I, *Streptococcus pneumoniae* polymerase, *Thermus aquaticus* polymerase, *Thermus flavus* polymerase, *Thermus thermophilus* polymerase, *Deinococcus radiodurans* polymerase, *Bacillus caldotenax* polymerase, *E. coli* bacteriophage T5 polymerase, *Thermatoga maritima* polymerase, mycobacteriophage L5 polymerase, and *E. coli* bacteriophage SP01 polymerase, wherein said mutant DNA polymerase comprises a substitution of Tyr for Phe at a position in said polymerase corresponding to Phe<sub>570</sub> of wild-type T5 polymerase.

8. A DNA molecule as claimed in claim 1, wherein said mutant protein is a mutant T5 DNA polymerase comprising a substitution of Tyr for Phe<sub>570</sub> of wild-type T5 polymerase.

9. The DNA molecule of claim 8, further comprising a promoter, wherein said promoter is in a position and orientation with respect to the coding

sequence such that the mutant protein may be expressed in a cell under the control of said promoter.

10. The molecule of claim 8, wherein said coding sequence is heterologous to the promoter.

11. A host cell comprising the DNA molecule of claim 8.

12. The host cell of claim 11, wherein said host cell is *E. coli*.

13. A method for producing a protein, wherein said protein is a mutant T5 DNA polymerase comprising a substitution of Tyr for Phe<sub>570</sub> of wild-type T5 polymerase, said method comprising:

- (a) culturing a host cell comprising the DNA molecule of claim 9, and
- (b) isolating said protein from said host cell.

14. A mutant DNA polymerase as claimed in claim 7, wherein said mutant DNA polymerase is a mutant T5 DNA polymerase comprising a substitution of Tyr for Phe<sub>570</sub> of wild-type T5 DNA polymerase.

15. A DNA molecule as claimed in claim 1, wherein said mutant protein is a mutant Taq DNA polymerase comprising a substitution of Tyr for Phe<sub>667</sub> of wild-type Taq polymerase.

16. The DNA molecule of claim 15, further comprising a promoter, wherein said promoter is in a position and orientation with respect to the coding sequence such that the mutant protein may be expressed in a cell under the control of said promoter.

17. The molecule of claim 16, wherein said coding sequence is heterologous to the promoter.

18. A host cell comprising the DNA molecule of claim 15.

19. The host cell of claim 18, wherein said host cell is *E. coli*.

20. A method for producing a protein, wherein said protein is a mutant Taq DNA polymerase comprising a substitution of Tyr for Phe<sub>667</sub> of wild-type Taq polymerase, said method comprising:

- (a) culturing a host cell comprising the DNA molecule of claim 16, and
- (b) isolating said protein from said host cell.

21. A mutant DNA polymerase as claimed in claim 7, wherein said mutant DNA polymerase is a mutant Taq DNA polymerase comprising a substitution of Tyr for Phe<sub>667</sub> of wild-type Taq DNA polymerase.

22. A DNA molecule as claimed in claim 1, wherein said mutant protein is a mutant Klenow fragment of *E. coli* DNA polymerase I comprising a substitution of Tyr for Phe<sub>762</sub> of wild-type Klenow fragment DNA polymerase.

23. The DNA molecule of claim 22, further comprising a promoter, wherein said promoter is in a position and orientation with respect to the coding sequence such that the mutant protein may be expressed in a cell under the control of said promoter.

24. The molecule of claim 23, wherein said coding sequence is heterologous to the promoter.

25. A host cell comprising the DNA molecule of claim 22.

26. The host cell of claim 25, wherein said host cell is *E. coli*.

27. A method for producing a protein, wherein said protein is a mutant Klenow fragment of *E. coli* DNA polymerase I comprising a substitution of Tyr for Phe<sub>762</sub> of wild-type Klenow fragment of *E. coli* DNA polymerase I, said method comprising:

(a) culturing a host cell comprising the DNA molecule of claim 23, and

(b) isolating said protein from said host cell.

28. A mutant DNA polymerase as claimed in claim 7, wherein said mutant DNA polymerase is a mutant Klenow fragment of *E. coli* DNA polymerase I comprising a substitution of Tyr for Phe<sub>762</sub> of wild-type Klenow fragment of *E. coli* DNA polymerase I.

09558421.042600

5

10

## **Abstract of the Invention**

### **Mutant DNA Polymerases and the Use Thereof**

5 The present invention relates to mutant DNA polymerases which incorporate dideoxynucleotides with about the same efficiency as deoxynucleotides. The present invention also related to mutant DNA polymerases which also have substantially reduced 5'-to-3' exonuclease activity or 3'-to-5' exonuclease activity. The invention also relates to DNA molecules coding for the mutant DNA polymerases, and hosts containing the DNA molecules.

p:\users\bobm\A4-03.WPD

009240 " 042600

00558421 042500

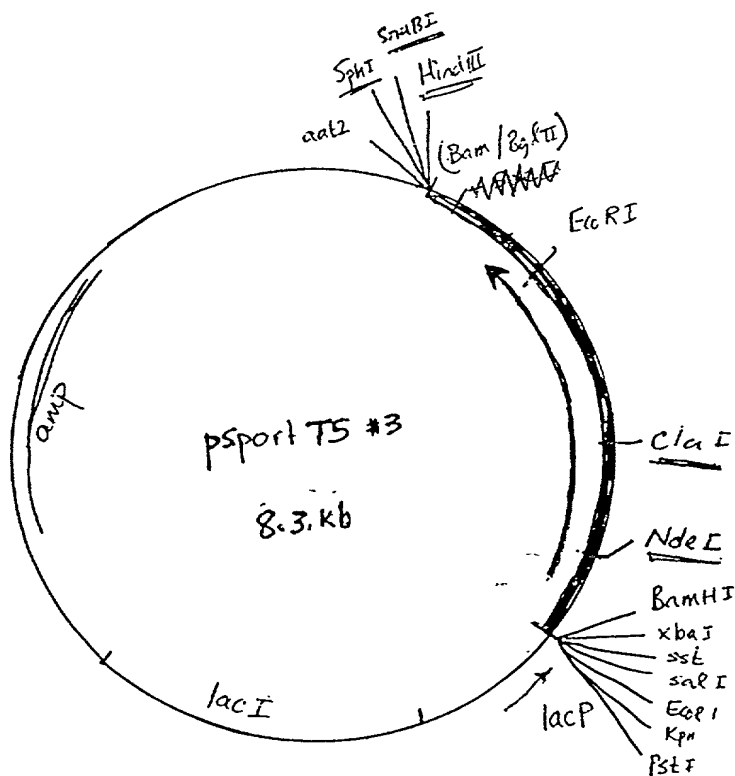


Fig. 1



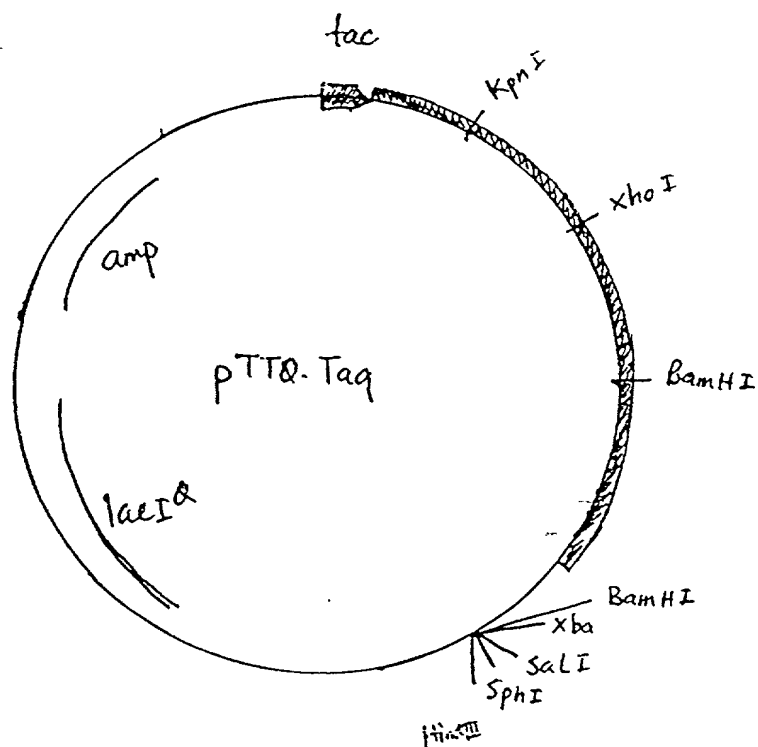
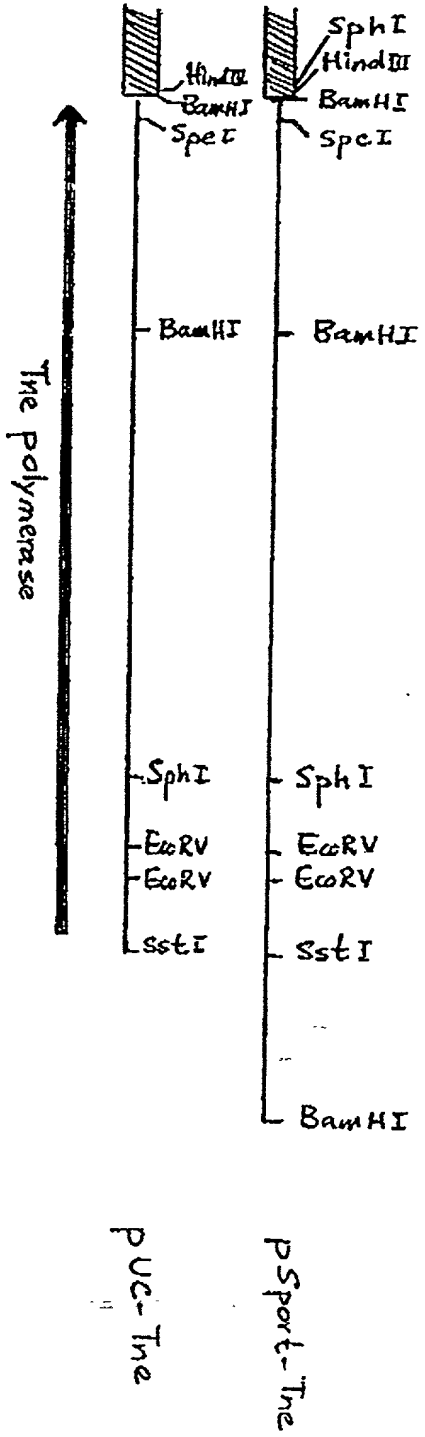


Fig. 2

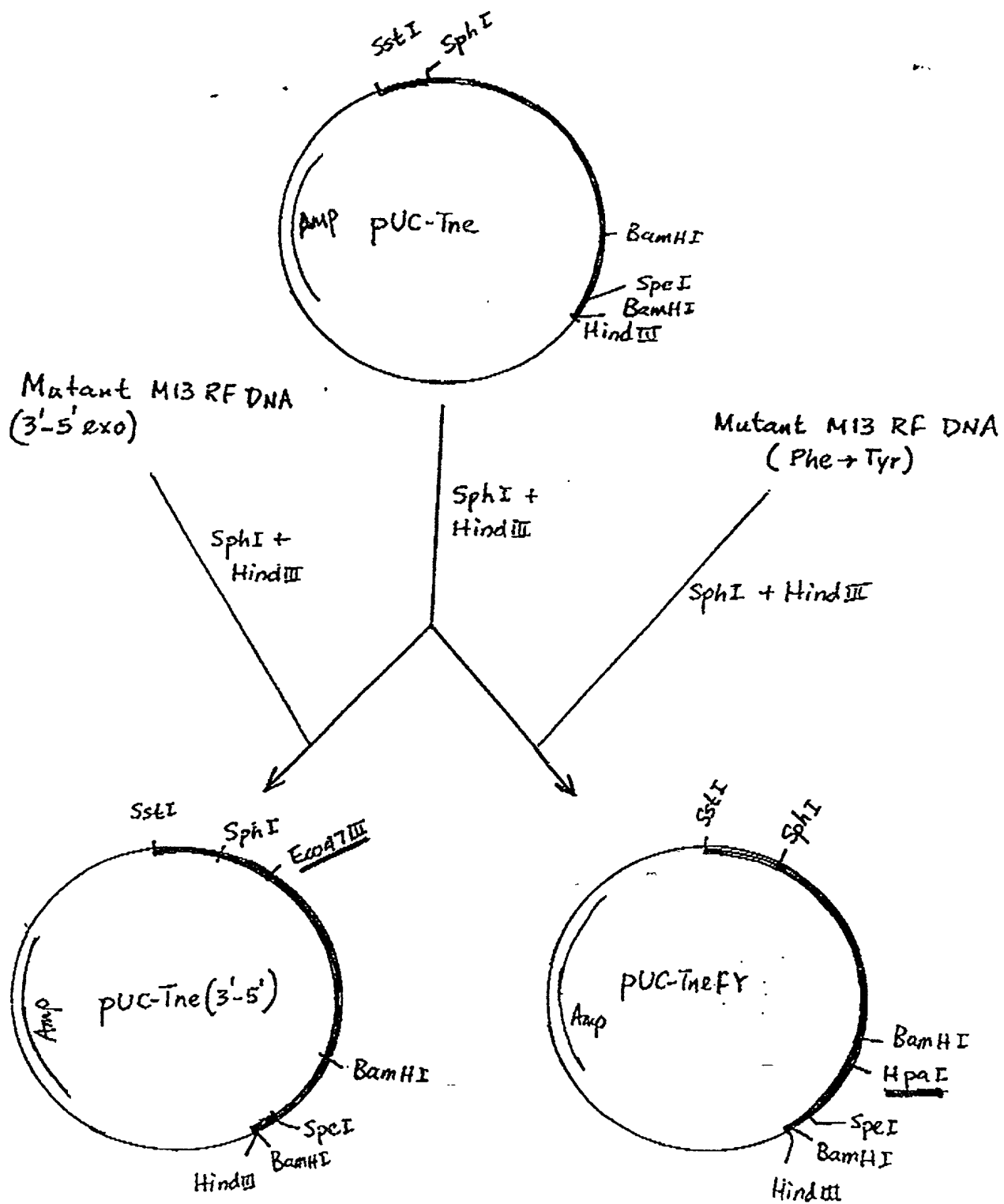
Fig. 3



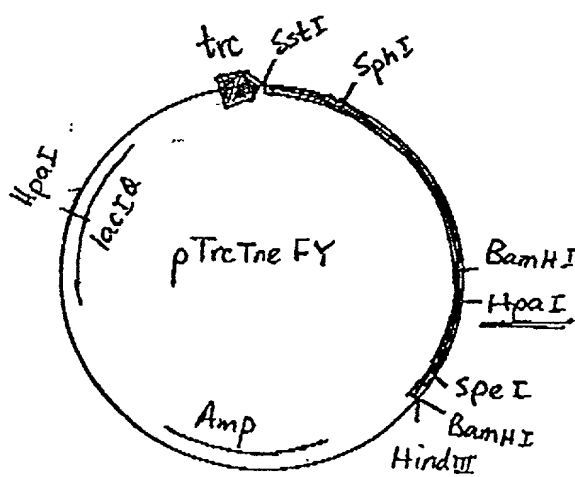
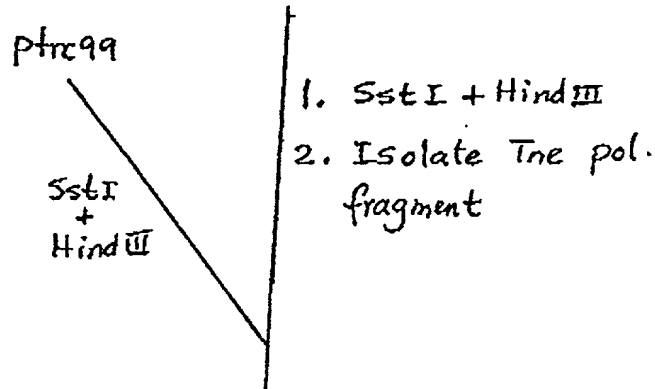
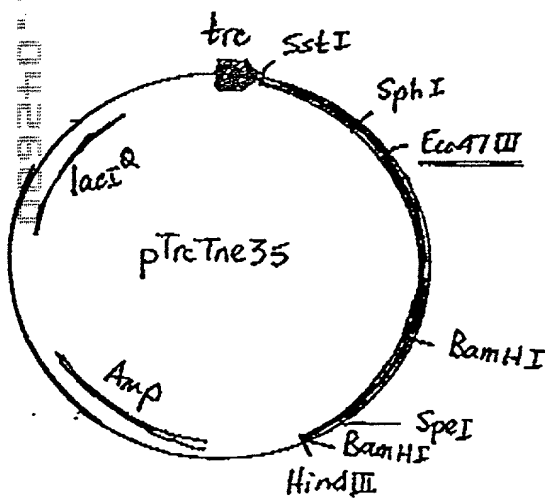
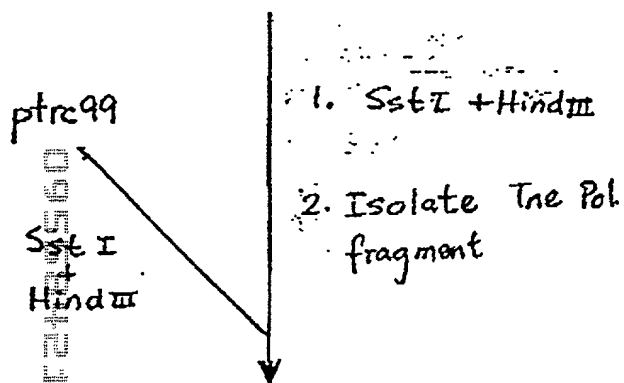
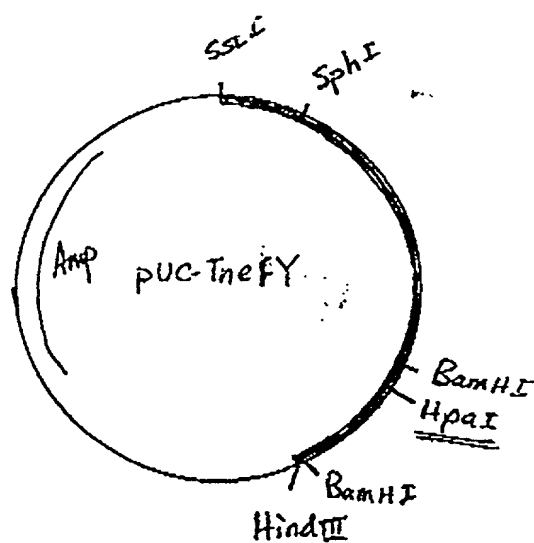
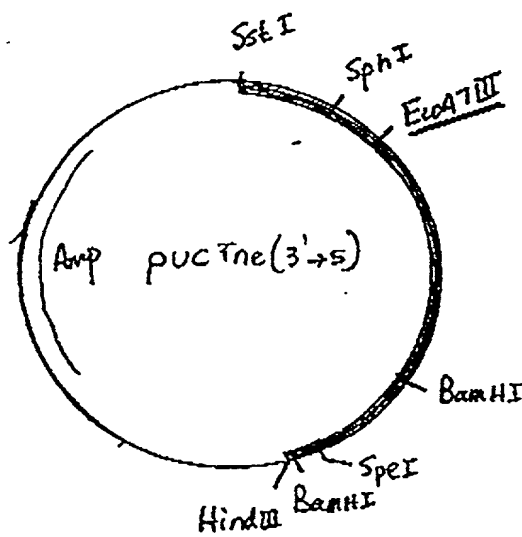
The region contains the O-helix homologous sequences.

09553421.042600

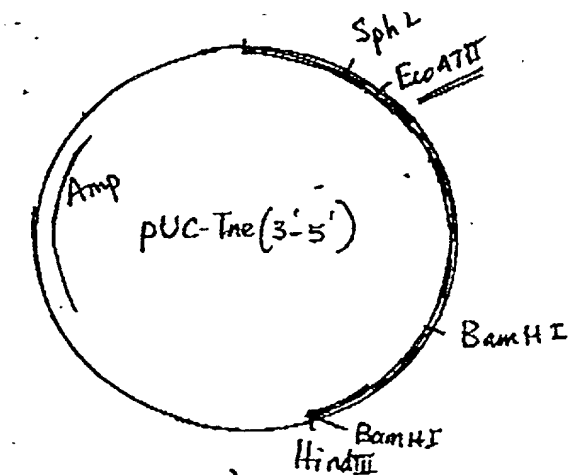
09558421.042600



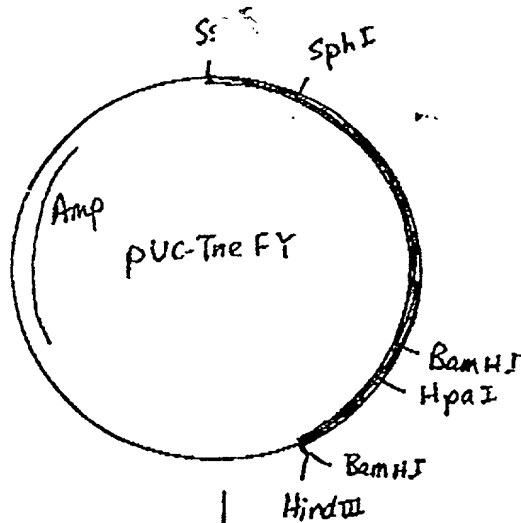
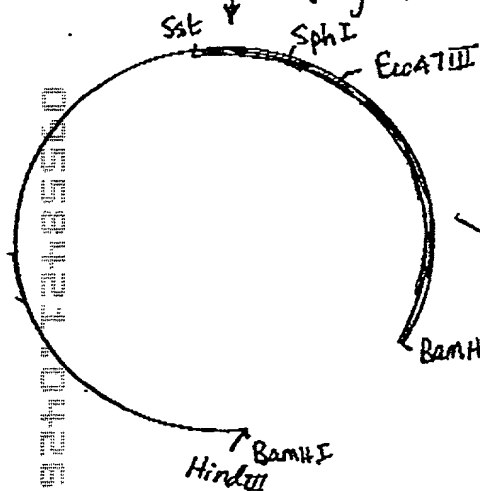
5A  
Fig. 3A



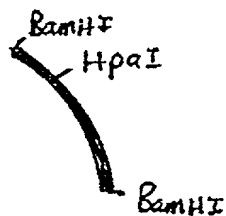
SB  
Fig 3B



1. Digest with BamHI
2. Dephosphorylate
3. Isolate the largest fragment



1. Digest with BamHI
2. Isolate the smallest fragment



ligate

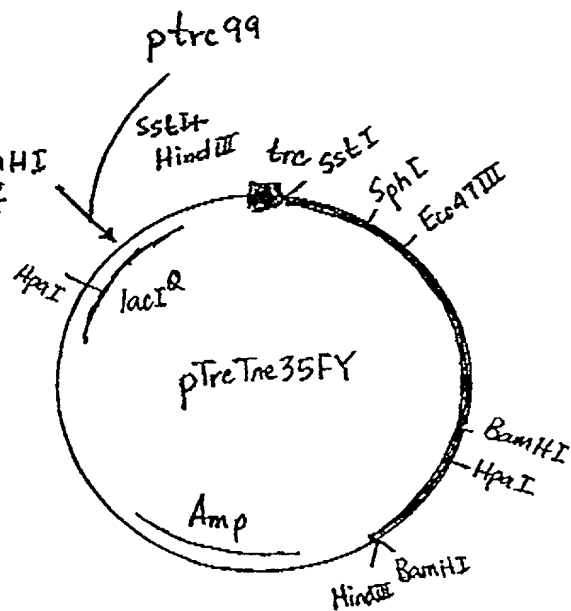
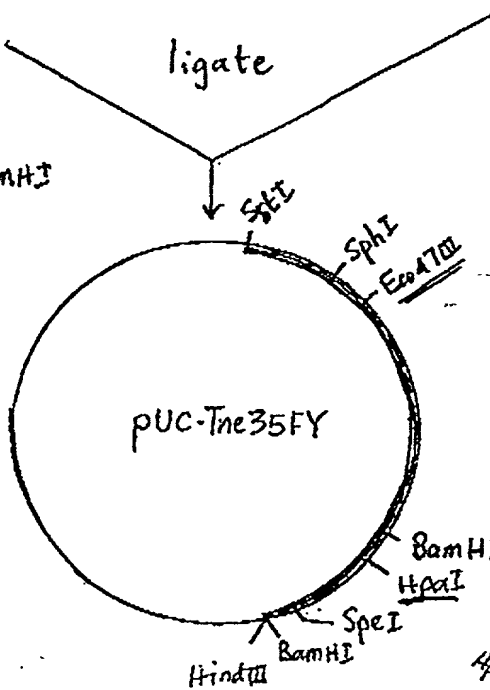


Fig. A<sup>6</sup>

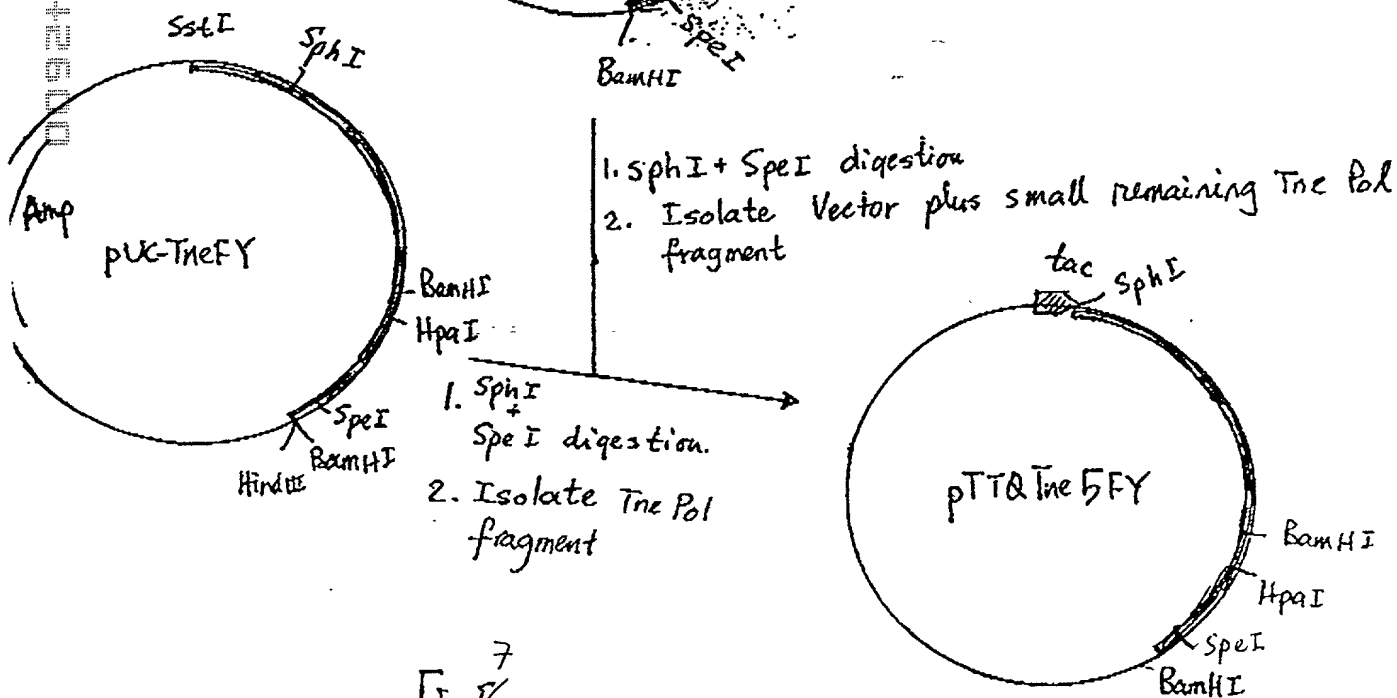
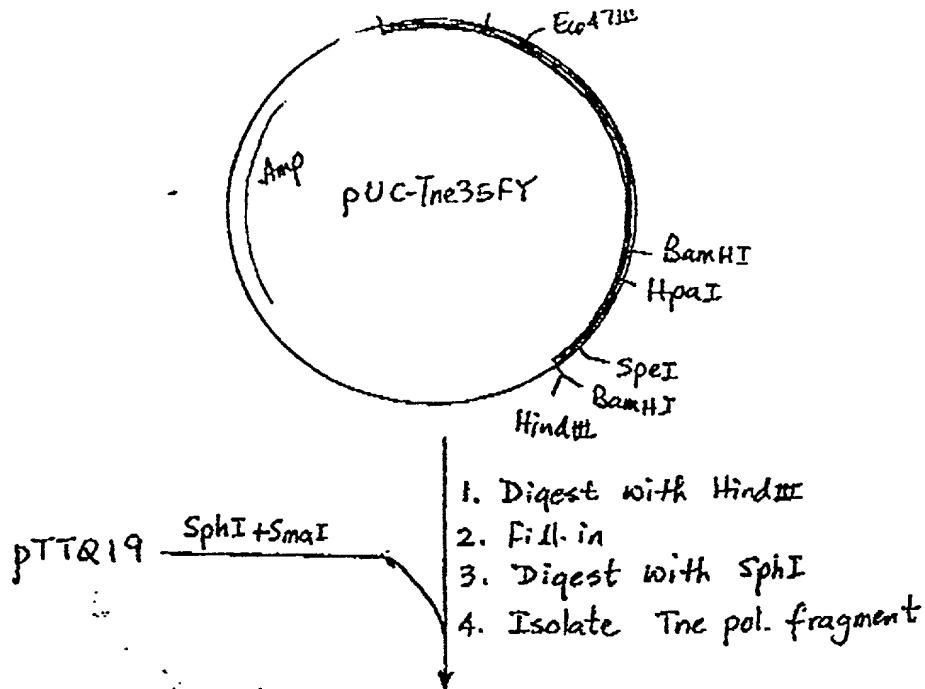


Fig. 5

## Declaration for Patent Application

Docket Number: 0942.3600002

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled Mutant DNA Polymerases and Uses Thereof, the specification of which is attached hereto unless the following box is checked:

- ☐ was filed on \_\_\_\_\_;  
as United States Application Number or PCT International Application Number \_\_\_\_\_; and  
was amended on \_\_\_\_\_ (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information that is material to patentability as defined in 37 C.F.R. § 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. § 119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate, or PCT International application having a filing date before that of the application on which priority is claimed.

### Prior Foreign Application(s)

Priority Claimed

_____ (Application No.)	_____ (Country)	_____ (Day/Month/Year Filed)	<input type="checkbox"/> Yes <input type="checkbox"/> No
_____ (Application No.)	_____ (Country)	_____ (Day/Month/Year Filed)	<input type="checkbox"/> Yes <input type="checkbox"/> No

I hereby claim the benefit under 35 U.S.C. § 119(e) of any United States provisional application(s) listed below.

_____ (Application No.)	_____ (Filing Date)
_____ (Application No.)	_____ (Filing Date)

I hereby claim the benefit under 35 U.S.C. § 120 of any United States application(s), or § 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. § 112, I acknowledge the duty to disclose information that is material to patentability as defined in 37 C.F.R. § 1.56 that became available between the filing date of the prior application and the national or PCT International filing date of this application.

08/537.397 (Application No.)	October 2, 1995 (Filing Date)	pending (Status - patented, pending, abandoned)
08/525.057 (Application No.)	September 8, 1995 (Filing Date)	pending (Status - patented, pending, abandoned)

Send Correspondence to:

STERNE, KESSLER, GOLDSTEIN & FOX  
1100 New York Avenue, N.W.  
Suite 600  
Washington, D.C. 20005-3934

Direct Telephone Calls to:

(202) 371-2600

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first inventor Deb K. CHATTERJEE	<i>Deb K. Chatterjee</i>
Inventor's signature	<i>Deb K. Chatterjee</i> 12/8/95 Date
Residence 6 Forest Ridge Court, N. Potomac, Maryland 20878	
Citizenship United States	
Post Office Address Same as above.	
Full name of second inventor	
Second Inventor's signature	Date
Residence	
Citizenship	
Post Office Address	



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Chatterjee, Deb K.
- (ii) TITLE OF INVENTION: Mutant DNA Polymerases and Uses Thereof
- (iii) NUMBER OF SEQUENCES: 43
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
  - (B) STREET: 1100 New York Avenue, N.W., Suite 600
  - (C) CITY: Washington
  - (D) STATE: DC
  - (E) COUNTRY: USA
  - (F) ZIP: 20005
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER: (To Be Assigned)
  - (B) FILING DATE: (Herewith)
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: US 08/525,057
  - (B) FILING DATE: 08-SEP-1995
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: US 08/537,397
  - (B) FILING DATE: 02-OCT-1995
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Esmond, Robert W.
  - (B) REGISTRATION NUMBER: 32,893
  - (C) REFERENCE/DOCKET NUMBER: 0942.3600002
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: 202-371-2600
  - (B) TELEFAX: 202-371-2540

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 40 base pairs

0955421 042600

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CAGGATCCAC ATGGTGCTTA ACGGCGACAT CCACACTAAG

40

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: both
  - (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTAACTTCT TGTGCGGTCT CAATGAC

27

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 11 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ala	Lys	Ala	Ile	Thr	Phe	Gly	Ile	Leu	Tyr	Gly
1				5					10	

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 11 amino acids

00558421 042600

(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

Ala Lys Thr Phe Ile Tyr Gly Phe Leu Tyr Gly  
1                 5                 10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

(2) INFORMATION FOR SEQ ID NO:6:

(ii) MOLECULE TYPE: protein

Ala Lys Ala Ile Asn Phe Gly Leu Ile Tyr Gly  
1 5 10

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Thr Phe Ile Tyr  
1

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ala Ile Thr Phe  
1

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Thr Ile Asn Phe  
1

0055421 04360  
0055421 04360

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 45 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: both  
    (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TCAGGCTGCT AAAACATTCA TCTACGGTAT ACTGTATGGT TCTGG

45

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 33 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: both  
    (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GTAGAGGACC CCGTAATTAA TGGTCTTGGC CGC

33

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 11 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gly Lys Met Val Asn Phe Ser Ile Ile Tyr Gly  
1                    5                    10

00958421 042600 "Feb 85 60"

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Val Asn Phe  
1

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 694 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 2..694

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

G	GAT	CCA	GAC	TGG	TGG	ATC	GTC	AGT	GCG	GAT	TAT	TCC	CAA	ATA	GAA	46
	Asp	Pro	Asp	Trp	Trp	Ile	Val	Ser	Ala	Asp	Tyr	Ser	Gln	Ile	Glu	
1				5					10						15	
CTC	AGA	ATC	CTC	GCT	CAT	CTC	AGT	GGT	GAT	GAG	AAC	CTT	GTG	AAG	GCC	94
Leu	Arg	Ile	Leu	Ala	His	Leu	Ser	Gly	Asp	Glu	Asn	Leu	Val	Lys	Ala	
			20						25				30			
TTC	GAG	GAG	GGC	ATC	GAT	GTG	CAC	ACC	TTG	ACT	GCC	TCC	AGG	ATC	TAC	142
Phe	Glu	Glu	Gly	Ile	Asp	Val	His	Thr	Leu	Thr	Ala	Ser	Arg	Ile	Tyr	
			35					40					45			
AAC	GTA	AAG	CCA	GAA	GAA	GTG	AAC	GAA	GAA	ATG	CGA	CGG	GTT	GGA	AAG	190
Asn	Val	Lys	Pro	Glu	Glu	Val	Asn	Glu	Glu	Met	Arg	Arg	Val	Gly	Lys	
			50					55					60			

0055421 043600

ATG GTG AAC TTC TCT ATA ATA TAC GGT GTC ACA CCG TAC GGT CTT TCT	238
Met Val Asn Phe Ser Ile Ile Tyr Gly Val Thr Pro Tyr Gly Leu Ser	
65 70 75	
GTG AGA CTT GGA ATA CCG GTT AAA GAA GCA GAA AAG ATG ATT ATC AGC	286
Val Arg Leu Gly Ile Pro Val Lys Glu Ala Glu Lys Met Ile Ile Ser	
80 85 90 95	
TAT TTC ACA CTG TAT CCA AAG GTG CGA AGC TAC ATC CAG CAG GTT GTT	334
Tyr Phe Thr Leu Tyr Pro Lys Val Arg Ser Tyr Ile Gln Gln Val Val	
100 105 110	
GCA GAG GCA AAA GAG AAG GGC TAC GTC AGG ACT CTC TTT GGA AGA AAA	382
Ala Glu Ala Lys Glu Lys Gly Tyr Val Arg Thr Leu Phe Gly Arg Lys	
115 120 125	
AGA GAT ATT CCC CAG CTC ATG GCA AGG GAC AAG AAC ACC CAG TCC GAA	430
Arg Asp Ile Pro Gln Leu Met Ala Arg Asp Lys Asn Thr Gln Ser Glu	
130 135 140	
GGC GAA AGA ATC GCA ATA AAC ACC CCC ATT CAG GGA ACT GCG GCA GAT	478
Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile Gln Gly Thr Ala Ala Asp	
145 150 155	
ATA ATA AAA TTG GCT ATG ATA GAT ATA GAC GAG GAG CTG AGA AAA AGA	526
Ile Ile Lys Leu Ala Met Ile Asp Ile Asp Glu Glu Leu Arg Lys Arg	
160 165 170 175	
AAC ATG AAA TCC AGA ATG ATC ATT CAG GTT CAT GAC GAA CTG GTC TTC	574
Asn Met Lys Ser Arg Met Ile Ile Gln Val His Asp Glu Leu Val Phe	
180 185 190	
GAG GTT CCC GAT GAG GAA AAA GAA GAA CTA GTT GAT CTG GTG AAG AAC	622
Glu Val Pro Asp Glu Glu Lys Glu Glu Leu Val Asp Leu Val Lys Asn	
195 200 205	
AAA ATG ACA AAT GTG GTG AAA CTC TCT GTG CCT CTT GAG GTT GAC ATA	670
Lys Met Thr Asn Val Val Lys Leu Ser Val Pro Leu Glu Val Asp Ile	
210 215 220	
AGC ATC GGA AAA AGC TGG TCT TGA	694
Ser Ile Gly Lys Ser Trp Ser	
225 230	

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 230 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Asp Pro Asp Trp Trp Ile Val Ser Ala Asp Tyr Ser Gln Ile Glu Leu  
1 5 10 15  
Arg Ile Leu Ala His Leu Ser Gly Asp Glu Asn Leu Val Lys Ala Phe  
20 25 30  
Glu Glu Gly Ile Asp Val His Thr Leu Thr Ala Ser Arg Ile Tyr Asn  
35 40 45  
Val Lys Pro Glu Glu Val Asn Glu Glu Met Arg Arg Val Gly Lys Met  
50 55 60  
Val Asn Phe Ser Ile Ile Tyr Gly Val Thr Pro Tyr Gly Leu Ser Val  
65 70 75 80  
Arg Leu Gly Ile Pro Val Lys Glu Ala Glu Lys Met Ile Ile Ser Tyr  
85 90 95  
Phe Thr Leu Tyr Pro Lys Val Arg Ser Tyr Ile Gln Gln Val Val Ala  
100 105 110  
Glu Ala Lys Glu Lys Gly Tyr Val Arg Thr Leu Phe Gly Arg Lys Arg  
115 120 125  
Asp Ile Pro Gln Leu Met Ala Arg Asp Lys Asn Thr Gln Ser Glu Gly  
130 135 140  
Glu Arg Ile Ala Ile Asn Thr Pro Ile Gln Gly Thr Ala Ala Asp Ile  
145 150 155 160  
Ile Lys Leu Ala Met Ile Asp Ile Asp Glu Glu Leu Arg Lys Arg Asn  
165 170 175  
Met Lys Ser Arg Met Ile Ile Gln Val His Asp Glu Leu Val Phe Glu  
180 185 190  
Val Pro Asp Glu Glu Lys Glu Glu Leu Val Asp Leu Val Lys Asn Lys  
195 200 205  
Met Thr Asn Val Val Lys Leu Ser Val Pro Leu Glu Val Asp Ile Ser  
210 215 220  
Ile Gly Lys Ser Trp Ser  
225 230

(2) INFORMATION FOR SEQ ID NO:16:

009240 042650





- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Gly Ser Arg Leu Val Asp Arg Gln Cys Gly Leu Phe Pro Asn Arg Thr  
 1 5 10 15  
  
 Gln Asn Pro Arg Ser Ser Gln Trp  
 20

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 60 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Glu Pro Cys Glu Gly Leu Arg Gly Gly His Arg Cys Ala His Leu Asp  
 1 5 10 15  
  
 Cys Leu Gln Asp Leu Gln Arg Lys Ala Arg Arg Ser Glu Arg Arg Asn  
 20 25 30  
  
 Ala Thr Gly Trp Lys Asp Gly Glu Leu Leu Tyr Asn Ile Arg Cys His  
 35 40 45  
  
 Thr Val Arg Ser Phe Cys Glu Thr Trp Asn Thr Gly  
 50 55 60

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 99 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single

009558421.042500

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Arg Ser Arg Lys Asp Asp Tyr Gln Leu Phe His Thr Val Ser Lys Gly  
1 5 10 15  
Ala Lys Leu His Pro Ala Gly Cys Cys Arg Gly Lys Arg Glu Gly Leu  
20 25 30  
Arg Gln Asp Ser Leu Trp Lys Lys Lys Arg Tyr Ser Pro Ala His Gly  
35 40 45  
Lys Gly Gln Glu His Pro Val Arg Arg Arg Lys Asn Arg Asn Lys His  
50 55 60  
Pro His Ser Gly Asn Cys Gly Arg Tyr Asn Lys Ile Gly Tyr Asp Arg  
65 70 75 80  
Tyr Arg Arg Gly Ala Glu Lys Lys Lys His Glu Ile Gln Asn Asp His  
85 90 95  
Ser Gly Ser

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 8 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Arg Thr Gly Leu Arg Gly Ser Arg  
1 5

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 amino acids

09558421.042600

- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Gly Lys Arg Arg Thr Ser  
1 5

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Ser Gly Glu Glu Gln Asn Asp Lys Cys Gly Glu Thr Leu Cys Ala Ser  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 9 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

His Lys His Arg Lys Lys Leu Val Leu  
1 5

(2) INFORMATION FOR SEQ ID NO:26:

00938424 "042600

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 13 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Ile	Gln	Thr	Gly	Gly	Ser	Ser	Val	Arg	Ile	Ile	Pro	Lys
1				5					10			

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 14 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Asn	Ser	Glu	Ser	Ser	Leu	Ile	Ser	Val	Val	Met	Arg	Thr	Leu
1				5					10				

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 11 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Arg	Pro	Ser	Arg	Arg	Ala	Ser	Met	Cys	Thr	Pro
1				5					10	

009240 "T248550

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Leu Pro Pro Gly Ser Thr Thr  
1 5

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 4 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Ser Gln Lys Lys  
1

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Thr Lys Lys Cys Asp Gly Leu Glu Arg Trp  
1 5 10

009240" 12435550

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 10 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Tyr Thr Val Ser His Arg Thr Val Phe Leu  
1                    5                    10

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 11 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Asp Leu Glu Tyr Arg Leu Lys Lys Gln Lys Arg  
1                    5                    10

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 56 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Leu Ser Ala Ile Ser His Gly Ile Gln Arg Cys Glu Ala Thr Ser Ser

09558421.042600

Climatic Data		Soil Data		Vegetation Data		Water Data		Air Data	
Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value
Temperature (°C)	25.5	pH	7.2	Plant Height (cm)	150	Water Level (m)	1.2	Air Quality (ppm)	120
Humidity (%)	65	Moisture (%)	45	Leaf Area (cm²)	120	Water Flow (L/min)	0.8	CO2 Concentration (ppm)	400
Light Intensity (lux)	1500	Salinity (g/L)	0.5	Stem Diameter (cm)	2.5	Water Temperature (°C)	20	Oxygen Saturation (%)	95
Wind Speed (m/s)	2.5	Soil Depth (cm)	10	Root Length (cm)	30	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Temperature (°C)	22	Soil Type	Clay	Chlorophyll Content	0.8	Water Quality (ppm)	100	Barometric Pressure (hPa)	1013
Soil Moisture (%)	40	Soil Fertility	High	Flowering Time (days)	15	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil pH	7.5	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low	Soil Fertility	High	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Fertility	High	Soil Salinity	Low	Seed Germination (%)	90	Water Quality (ppm)	100	Relative Humidity (%)	70
Soil Salinity	Low								

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 19 amino acids  
 (B) TYPE: amino acid



(C) STRANDEDNESS: single  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Ser	Phe	Arg	Phe	Met	Thr	Asn	Trp	Ser	Ser	Arg	Phe	Pro	Met	Arg	Lys
1				5				10						15	

Lys Lys Asn

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 9 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Asn	Ser	Leu	Cys	Leu	Leu	Arg	Leu	Thr
1			5					

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 7 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Ala	Ser	Glu	Lys	Ala	Gly	Leu
1			5			

009940" 243550

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 11 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Pro	Ser	Phe	Ala	Leu	Asp	Leu	Glu	Thr	Ser	Ser
1				5					10	

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 11 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Pro	Tyr	Phe	Ala	Phe	Asp	Thr	Glu	Thr	Asp	Ser
1				5					10	

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 11 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Gly	Pro	Val	Ala	Phe	Asp	Ser	Glu	Thr	Ser	Ala
1				5					10	

09558421.042600

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 10 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Met	Ile	Val	Ser	Asp	Ile	Glu	Ala	Asn	Ala
1				5				10	

009240 "T218560

## \* \* \* S E Q U E N C E \* \* \*

BamHI

1 GGATCCAGAC TGGTGGATCG TCAGTGGCGA TTATTCCCAA ATAGAACTCA GAATCCTCGC  
G S R L V D R Q C G L F P N R T Q N P R  
→ D P D W W I V S A D Y S Q I E L R I L  
I Q T G G S S V R I I P K - N S E S S

61 TCATCTCAGT GGTGATGAGA ACCTTGTAAC GGCCTTCGAG GAGGGCATCG ATGTGCACAC  
S S Q W - - E P C E G L R G G H R C A H  
→ A H L S G D E N L V K A F E E G I D V H  
L I S V V M R T L - R P S R R A S M C T

121 CTTGACTGCC TCCAGGATCT ACAACGTAAA GCCAGAAGAA GTCAACCAAG AAATGCGACG  
L D C L Q D L Q R K A R R S E R R N A T  
→ T L T A S R I Y N V K P E E V N E E M R  
P - L P P G S T T - S Q K K - T K K C D

181 GGTGGAAG ATGGTGAAC TCTCTATAAT ATACGGTGTC ACACCGTACG GTCTTTCTGT  
G W K D G E L L Y N I R C H T V R S F C  
→ R V G K M V N (F) S I I Y G V T P Y G L S  
G L E R W - T S L - Y T V S H R T V F L

241 GAGACTTGCA ATACCGGTTA AAGAAGCAGA AAAGATGATT ATCAGCTATT TCACACTGTA  
E T W N T G - R S R K D D Y Q L F H T V  
→ V R L G I P V K E A E K M I I S Y F T L  
- D L E Y R L K K Q K R - L S A I S H C

301 TCCAAAGGTG CGAAGCTACA TCCAGCAGGT TGTTCAGAG GCAAAAGAGA AGGGCTACGT  
S K G A K L H P A G C C R G K R E G L R  
→ Y P K V R S Y I Q Q V V A E A K E K G Y  
I Q R C E A T S S R L L Q R Q K R R A T

361 CAGGACTCTC TTTGGAAGAA AAAGAGATAT TCCCCAGCTC ATGGCAAGGG ACAAGAACAC  
Q D S L W K K K R Y S P A H G K G Q E H  
→ V R T L F G R K R D I P Q L M A R D K N  
S G L S L E E K E I F P S S W Q G T R T

421 CCAGTCCGAA GGGGAAAGAA TCGCAATAAA CACCCCCATT CAGGGAACGT CGGCAGATAT  
P V R R R K N R N K H P H S G N C G R Y  
→ T Q S E G E R I A I N T P I Q G T A A D  
P S P K A K E S Q - T P P F R E L R Q I

481 AATAAAATTG GCTATGATAG ATATAGACGA GCAGCTGACA AAAAGAAACA TGAATCCAG  
N K I G Y D R Y R R G A E K K K H E I Q  
→ I I K L A M I D I D E E L R K R N M K S  
- - N W L - - I - T R S - E K E T - N P

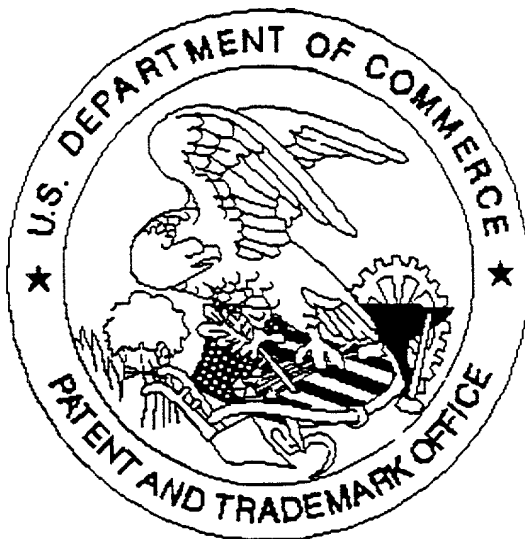
541 AATGATCATT CAGGTTTCATG ACGAACTGGT CTTCGAGGTT CCCGATGAGG AAAAAGAAGA  
N D H S G S - R T G L R G S R - G K R R  
→ R M I I Q V H D E L V F E V P D E E K E  
E - S F R F M T N W S S R F P M R K K K

501 ACTAGTTGAT CTGGTGAAGA ACAAATGAC AAATGTGGTG AAATCTCTCTG TGCCTCTTGA  
T S - S G E E Q N D K C G E T L C A S -  
→ E L V D L V K N K M T N V V K L S V P L  
N - L I W - R T K - Q M W - N S L C L L

661 GGTGACATA AGCATCGGAA AAAG GTG TTGA  
G - H K H R K K L V L  
→ E V D I S I G K S W S -  
R L T - A S E. K A G L

009240 1248560

United States Patent & Trademark Office  
Office of Initial Patent Examination -- Scanning Division



Application deficiencies were found during scanning:

☐ Page(s) figure 4 of the drawings were not present  
for scanning. (Document title)

☐ Page(s) \_\_\_\_\_ of \_\_\_\_\_ were not present  
for scanning. (Document title)

*There are only 7 sheets of drawings, not 9.*

☐ Scanned copy is best available.